



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 117364

To: **Phuong Bui**
Location: **rem 2a15 & 2c18**
Art Unit: **1638**
Tuesday, March 23, 2004

Case Serial Number: **09/913064**

From: **Beverly Shears**
Location: **Remsen Bldg.**
RM 1A54
Phone: **571-272-2528**

beverly.shears@uspto.gov

Search Notes

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:29:48 ; Search time 5324 Seconds
(without alignments)
10681.078 Million cell updates/sec

Title: US-09-913-064a-13
Perfect score: 1312
Sequence: 1 Gcagcagcactctctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.in.v.*
- 32: em.htg.other.*
- 33: em.htg.mu.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.vrt.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mu.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	833.2	63.5	1216	8	FSU31544	U31544 Pisum sativ
2	673.2	51.3	1474	8	CTAJ5081	AJ005081 Cyamopsi
3	649.2	49.5	1170	8	BT000032	BT000032 Arabidops
4	649.2	49.5	1313	8	AY054207	AY054207 Arabidops
5	649.2	49.5	1420	8	AY120709	AY120709 Arabidops
6	644.4	49.1	1056	8	AX507662	AX507662 Sequence
7	644.4	49.1	1056	8	AF334724	AF334724 Arabidops
8	642.8	49.0	1356	8	ATUDGE	Z54244 A.thaliana
9	631	48.1	1263	8	AK117913	AK117913 Arabidops
10	629.4	48.0	1397	8	AY085887	AY085887 Arabidops
11	621.6	47.4	1240	8	AY094481	AY094481 Arabidops
12	482.8	36.8	1375	8	AK104977	AK104977 Oryza sat
13	482.8	36.8	1425	8	AK073610	AK073610 Oryza sat
14	479.6	36.6	1223	8	AB097460	AB097460 Oryza sat
15	476.4	36.3	1516	8	BT009044	BT009044 Triticum
16	456.4	34.8	1360	8	CTAJ5082	AJ005082 Cyamopsi
17	445	33.9	1065	6	AX653074	AX653074 Sequence
18	445	33.9	1346	8	AB087745	AB087745 Oryza sat
19	435	33.2	1735	8	AB096863	AB096863 Oryza sat
20	435	33.2	1878	8	AK066269	AK066269 Oryza sat
21	433.4	33.0	1420	8	AF303682	AF303682 Zea mays
22	428.4	32.7	1286	8	AY085528	AY085528 Arabidops
23	423.6	32.3	1087	8	AY117180	AY117180 Arabidops
24	423.6	32.3	1356	8	AY065354	AY065354 Arabidops
25	423.6	32.3	2800	8	AY140073	AY140073 Arabidops
26	416.8	31.8	1316	8	AY084615	AY084615 Arabidops
27	416.4	31.7	1419	8	AK121426	AK121426 Oryza sat
28	416	31.7	1084	8	BT008539	BT008539 Arabidops
29	416	31.7	1393	8	AK118722	AK118722 Arabidops
30	385.6	29.4	1335	6	AX653761	AX653761 Sequence
31	313.6	23.9	10398	1	AE015612	AE015612 Shewanell
32	296.2	22.6	276034	1	AE017015	AE017015 Bacillus
33	287	21.9	17317	1	AY147913	AY147913 Streptoco
34	278.2	21.2	4626	14	AF065660	AF065660 Stealth v
35	277.8	21.2	1017	6	BD170587	BD170587 UDP-N-Ace
36	277.8	21.2	6715	1	BSGALE	X93339 B.subtilis
37	277.8	21.2	65143	1	D83026	D83026 Bacillus su
38	277.8	21.2	201139	1	BSUB0020	Z99123 Bacillus su
39	271.6	20.7	306050	1	BX321858	BX321858 Nitrosomo
40	270.6	20.6	250150	1	AP005342	AP005342 Vibrio vu
41	269.8	20.6	11351	1	AE004405	AE004405 Vibrio ch
42	267.6	20.4	1390	3	AK114505	AK114505 Ciona int
43	267.6	20.4	69727	6	AX770896	AX770896 Sequence
44	267.6	20.4	124056	1	BX571875	BX571875 Photornab
45	267	20.4	302000	1	AP003187	AP003187 Clostridi

ALIGNMENTS

RESULT 1	PSU31544	1216 bp	mRNA	linear	PLN 02-FEB-1996
LOCUS	Pisum sativum	UDP-galactose-4-epimerase (gale)	mRNA	complete cds.	
DEFINITION	U31544				
ACCESSION	U31544				
VERSION	U31544.1	GI:1173554			
KEYWORDS					
SOURCE	Pisum sativum (pea)				
ORGANISM	Pisum sativum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;				
	Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;				
	Pisum.				
REFERENCE	1 (bases 1 to 1216)				
AUTHORS	Slocum,R.D., Lake,M.R. and Williamson,C.L.				

TITLE Cloning and characterization of a UDP-galactose-4-epimerase ('galactowaldenase') and its expression in pea tissues

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1216)

AUTHORS Slocum, R.D.

TITLE Direct Submission

JOURNAL Submitted (12-JUL-1995) Robert D. Slocum, Biological Sciences, Goucher College, Baltimore, MD 21204-2794, USA

FEATURES

Location/Qualifiers

1..1216

/organism="Pisum sativum"

/mol_type="mRNA"

/cultivar="WandoPea"

/db_xref="taxon:3888"

1..1216

/gene="gale"

6..1058

/gene="gale"

/EC number="5.1.3.2"

/function="catabolism of galactose to glucose in Leloir pathway, and in galactose synthesis from glucose."

/note="galactowaldenase"

/codon_start=1

/product="UDP-galactose-4-epimerase"

/protein_id="AAA86532.1"

/db_xref="GI:117355"

/translation="WVASSQKILVTGSAGFTGHTVWLLNNGFNVSIIIDNFSVME
AVREVGVNSLSONLSTGLDKNDLEKLSKEDAVIHPAGLKVAGSEVNP
RRYFNNLVGTLNLEWAKENCKWFSSTATYGOPEKIPCTVEDFKLOANNPYGRT
KLFLEIAHDIOKASPEWRIVLIRYFNPVGHESKLGEDPRGFNNIMLPYQVAVG
RLPELVYGHDPYPTDGAIRDYIHYMDLADGHIARLKLFTSENIQMLTAYNLGVRG
SSVLEWAAFEKASGKKALKLCPRRPGDTEVYASTAKAEKELGWAKYGVSEWCRD
QNWNAKNPNWGYSGKP"

ORIGIN

Query Match 63.5%; Score 833.2; DB 8; Length 1216;
Best Local Similarity 84.0%; Pred. No. 1.7e-191;
Matches 953; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 31 AGCATGGTGCTTCCTCCCAACACATTCGTGGTCACCGGTGGTGGCGGTTTCATGGCAC 90

DB 3 ACATGGTGGCTTCGTGCGAAGATACCTGTAACCGGAGTGGCGGTTTCATGGCACT 62

QY 91 CACACGGTGGCTTCCTCAAGCTGGCTTCAGGTTTCATATCGACATTTCCGAT 150

DB 63 CACACGGTGGCTTCCTCAAGCTGGCTTCAGGTTTCATATCGACATTTCCGAT 122

QY 151 AACTCCGTGATGAAGCAGTGGACCGGTCGCGCAAGTGGTGGCCCTCTGCTTTCTCAG 210

DB 123 AATTCTGTTATGAGACAGTGGACCGGTCGCGCAAGTGGTGGCCCTCTGCTTTCTCAG 182

QY 211 AACTCCGATTCACCGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTCTCC 270

DB 183 AATCTCGAATTCATCTGGGAGATCTCAGGAATAAGATGATTTGGAGAACTCTCTCA 242

QY 271 AAACACATTTGATGCGGTATCCATTTGCTGGCTTGAAGCGGTGCTGAAGCGTT 330

DB 243 AATCTAAATTTGATGCTGATTCATCTTCTGGATGAAGCACTTGGTGGAGTGT 302

QY 331 GCGAAGCCCGTGGCTATTTGATTTTAAATTTGGTGGCAACATCACTCTACGAGTTT 390

DB 303 GAAATCTCTGCTGTTATTTTCGATAATAATCTTTGGAACATCAATCTCTATGAAGTT 362

QY 391 ATGCAAGATTAATTCGAAGATGTTTCTCATCTCGCAACCGTTTATGGCCAA 450

DB 363 ATGGCTTAAGCATTAATTTGAAGATGTTTCTCGTATCTGCAACTGTTTATGCCAA 422

QY 451 CCTGAAAGATACCGTGTGAGGAGATTTCAAGTTACAGCTATGATCCCTATGACCG 510

DB 423 CCTGAAAGATACCGTGTGAGGAGATTTCAAGTTACAGCTATGATCCCTATGACCG 482

QY 511 ACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAGCTGAACAGATGAAG 570

DB 483 ACCAAGCTTTCTCTTGAAGAAATTCACGAGATATTCAGAAAGCTGAGCCAGATGGAGA 542

QY 571 ATCATATTAATGAGATACCTTCAATCAGTTGGGGCTCATGAAGTGGCAAACTCGTGAA 630

DB 543 ATCGTTTACTGCGGTACTTCAATCCAGTCGGGGCACATGAAGTGGTAACTTGTGAA 602

QY 631 GATCCCAAGGGCATCCCAATAACCTTCATGCTTACATTCAGCAAGTACTGTGTGAAGA 690

DB 603 GATCCCAAGGGCATCCCAATAATCTCATGCTTATATACAGCAAGTACCGCTGGAAGA 662

QY 691 TTGACTGAATCAATGTATATACGGTCAATGATTATCCAAACAGGATGGCTTCGGATCCGG 750

DB 663 TTACCTGAGCTCAATGTATATGATGCTCATGATTATCCACAGGATGGCTTCGCAACGG 722

QY 751 GACTATATCCATGTGATGACTTGGCAGATGGCCATATTCGTCCTCGGAAAGCTCTTC 810

DB 723 GACTATATCCATGTGATGACTTGGCAGATGGCTCATATTCGTCCTCGGAAAGCTCTTC 782

QY 811 ACAACGGAGAACATAGTGTGATCTTACACCTGGGAACCTGGTCTGTGGAACATCTGTC 870

DB 783 ACATCAGAAACATTCGTTGATCTGCTTATAATTTGGGAACCTGGTCTGTGCTGTCG 842

QY 871 CTTGAAATGGTTACAGCATTTGAAAGGCTTCGGCAAGAAATTCAGTAAATATATGT 930

DB 843 CTTGAAATGGTTGCTGCAATTTGAGAAGCTTCGGCAGAAATTCGATGAATTTGTT 902

QY 931 CCAAGAGACCGGAGATCGACTGAGTTTATGATCTTACAGAGAGACTGAGAAAGAA 990

DB 903 CCAAGAGACCGGAGATCGACTGAGTTTATGATCTTACAGAGAGACTGAGAAAGAA 962

QY 991 CTTGTTTGGAGGCAACATGATGTTGGAGGATGTCGAGGACCAATCGAATTTGGGCA 1050

DB 963 CTTGTTTGGAGGCAACATGATGTTGGAGGATGTCGAGGACCAATCGAATTTGGGCA 1022

QY 1051 AAGAACATCCCTGGGTTACGGGGAGCCCTTGAATTAAGTTG---AGAAATATACTG 1107

DB 1023 AAGAACATCCCTGGGTTTACTCAGGAAGCTTGAATCAGTTGACATAAATACTA 1082

QY 1108 CTGATCTCAAGTCTTTTACATATAAGGCTCTCTTATAGATACTTTT 1161

DB 1083 CACATATGCCAATCTTTCTGTAAGATAGGCTCCCATTTGTTAAGATACTT 1136

RESULT 2

CTAJ5081 1474 bp mRNA linear PLN 27-APR-1999

LOCUS Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase, clone

DEFINITION

ACCESSION AJ005081

VERSION AJ005081.1 GI:3021354

KEYWORDS UDP-galactose 4-epimerase.

SOURCE Cyamopsis tetragonoloba (guar)

ORGANISM Cyamopsis tetragonoloba

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Indigoferae; Cyamopsis.

REFERENCE 1

AUTHORS Brunstedt, J., Joerbo, M., Pedersen, S.G. and Marcussen, J.

TITLE Isolation and expression of two cDNA clones encoding UDP-galactose epimerase expressed in developing seeds of the endospermous legume guar

JOURNAL Plant Sci. 142, 147-154 (1999)

REFERENCE 2 (bases 1 to 1474)

AUTHORS Brunstedt, J.

TITLE Direct Submission

JOURNAL Submitted (24-MAR-1998) Danisco Biotechnology, Danisco A/S, Langebrogade 1, PO Box 17, Copenhagen DK-1001, Denmark

FEATURES

Location/Qualifiers

1..1474

/organism="Cyamopsis tetragonoloba"

/mol_type="mRNA"

/db_xref="taxon:3832"

Qy	814	ACGAGAACATAGTGTGTTACTCTTCAACACTGGGAACCTGGTGGTGAACATCTGTGCTT	873
Db	871	ACAGACAACATCTGGTGTGTTACTCGCTCTAACTTGGGAACCTGGTGGTGTACGTGAGTGCCT	930
Qy	874	GAATGTGTTACAGCATTTGAAAAAGGCTTCGGCAAGAAAAATTCAGCTAAATATTGTCCA	933
Db	931	GAATGTGTTGCAGCATTTGAAAAAGCTTCGGGCAAGAAAAATTCATAAAAAATGTGTCCA	990
Qy	934	AGAAGACCGGGAGATCGGACTGAGGTTTATGCATCTACAGAGAGAGCTGAGAAGAACTT	993
Db	991	AGAGGCGCAGGGGATGCTACTGCTGTTTATGTCATCTACGGAGAGGCTGAGAAGAACTT	1050
Qy	994	GGTTGGAAGGCAAACTACTGTTGTGGAGGAGATGTGAGGACCAATGGAATTTGGGCAAG	1053
Db	1051	GGTTGGAAGGCAAAATACGCTGTGGAAGAAATGTGCAGGACCAATGGAATTTGGGCAAGC	1110
Qy	1054	AACATCCCTGGGGTTACGGGGGAGGCTTGAATTAGCTTGAGAAATAT	1103
Db	1111	AATATCCATGGGGGTATCAAGGGAAGCATGATTTTTTTTCTTTATAT	1160

RESULT 3

BT000032 1170 bp mRNA linear PLN 19-SEP-2002

LOCUS Arabidopsis thaliana uridine diphosphate glucose epimerase

DEFINITION (At1512780) mRNA, complete cds.

ACCESSION BT000032

VERSION BT000032.1 GI:23197647

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Spurmatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1170)

AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Sacou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Becker,J., Theologis,A. and Davis,R.W.

TITLE Direct Submission

JOURNAL Submitted (19-SEP-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT e-mail for correspondence: arabsequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M., Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W., Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Torum,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers

1. .1170

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="U16095"

/ecotype="Columbia"

FEATURES

source


```

/note="This clone is in pENTR/SD-dTropo. This is a cloned
PCR product using RIKEN clone RAF105-11-004 (AY120709) as
a template"
1. 1.1170
  /gene="At1g12780"
  /note="synonym: F13K23.3"
1. 1.1056
  /gene="At1g12780"
  /codon_start=1
  /product="uridine diphosphate glucose epimerase"
  /protein_id="AA15351.1"
  /db_xref="GI:23197648"
  /translation="MGSSVEQNILVTGGAGFTGTHVQLLKDGFKVISIINDFNDSVVI
EAYVRVNNIVGPDLSKLTFLNLDGKIDIEKLSKQRPDAVTHFAFLKAVGSEVSN
PRAYFNNILVGLNLTWYKNNKCMYFSSATVYGGPEKI PCWEDFELKAMPYGR
KLPLEIARDI OKAPENWKRI ILLRYFNFPVGAHSGISGDPKGI GNLLMPLYIQQAV
GRUPLNLYVGHDPYTEDGSAVRDYTHVMDLADGHLAALRLFADPKTGCTAYNLGTGQV
GLPFLVLEWAFAPGKASQKIKLCPPRRSDATAVVASTAEKAEGLMKVYGDVDMCR
DQKSWNNANWFGKQKSI"

```

gene

```
/gene="ALG12780"  
/note="synonym: F13K23.3"
```

1. 1056

```
/gene="A:
```

```
/codon_start=1
```

```
product="npod/
```

```
/protein_id="AA15351.1"
```

```
/ow_xlet="GI:23197848"
/translation="MGSSVEONTI
```

EAVDRVRELVGPDLSKKLDEN

PRRYPDNNLVGTINLYETWAKYCNCKMYFSSATVYSGQPKICMEDFELKAMNPYGV
 TKLFLBEIARDI QKABPEWRI LLRYFNPVGAHSGSGEDPKGINNLMLPILQVAV
 GRLPENLVGHDPYTEDGSAVRDIYHMDLADGHIAALRKLADPKI.GCTA.VNLGTQGV
 GTSVLEWMAAF EKASGKTKI PKLCPRRSGDATVAYASTAEKAEKELHKNKATGVDEMCRI
 DOWKANNNDPMGWGKTKI."

ORIGIN

Query Match	49.5%;	Score 649.2;	DB 8;	Length 1170;
Best Local Similarity	76.3%;	Pred. No. 8.7e-147;		
Matches 798;	Conservative	0;	Mismatches 248;	Indels 0;
Gaps 0;				

49 CAA CACATTTCTGGTTCACCGGTGGTCCGGTTTCATTGSCACCCACACCGTGGTTCAGCTT 108
19 CAGAACATTTCTGTACTTGGTGGTGGTGGTCTGGTCTTATCGGACGCACTACTGTTGTTCAACTT 78
109 CTCAAAGCTGGCTTCAGCGTTTCAATAATCGACAATTTGATATACTCGTCAATGGAAGCA 168
79 CTCAAAGATGGTTTAAAGTTTTCGATCATCGATAATTTTGATACTCTGTTTATCGAAGCT 138
169 GTGGACCGCGTCCGCAAGTGGTGGGCCCTCTGCTTTCTCAGAACTTCCAATTCACCCAG 228
139 GTTGATAGATTAGGAGCTTGTGGTCTCTGATCTCTCCAGAAGCTGCATTTCAATCTG 198
229 GCGCATCTCCGGAATPAGGGATGACCTTGAGAAACCTCTTCTCCAAAACAATTTTGATGCC 288
199 GGTGATCTTAAGAAAACAAAGGGGACATTCAGAAACTATTTCTCCAGCAGAGATTTTGATGCT 258
289 GTGATCCACTTTGCTGGCTTGAAAGCGGTGCTGAAAGCGTTGGAAAGCCCGCTCGGTAT 348
259 GTGATTCATTTTGGCGGTCTTAAAGCTGTGGGTGAGAGTGTGTTGAAAACCTTCGCGCTAC 318
349 TTTGATTTTAATTTGGTTGGACCACTCAACCTCTACGAGTTTATGGCAAGATATAATGTC 408
319 TTTGACAATACTTGGTTGGAACTCAATCTATATGAGACCATGTGCAAGATCAACTGTC 378
409 AAAAGATGGTTTTCTCATCATCTGCAACCGTTTTATGGCCAACTCTGAAAAGATACCGGTG 468
379 AAAATGATGGTGTGTTTTCATCTCTGCCATGTTTATGGACAACCTGAAAAGATTCATATC 438
469 GAGGAGGATTTCAAGTTACAAGCTATGAATCCCTATGGACGAGCAACCAAGCTTTTCTGTGAA 528
439 ATGGAGAGCTTTGAATTTAAGGCTATGAATCTCTATGTCGTACTAAGCTCTTTTCTTGAA 498
529 GAAATTCGCGGAGATATTTCAGAAAGCTGAAACGAGATCGGAAGATCATATTACTGAGATAC 588
499 GAAATAGCTAGAGATATTCAAAGGCAGAAACCGGAATGGAGAATTAATCTGCTGAGGTAC 558
589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCCAAGGGGATCCCA 648
559 TTCAATCCGTGAGAGACATGAGAGTGGCAGTATTGGTGAGGATCCAAAGGCATCCCC 618
649 AATAACCTCATGCCCTTACATTCAGCAAGTAGCTGTTTGAAGATTGACTGAACCTCAATGTA 708
619 AATAACCTCATGCCCTTACATCCAAAGTGGCGGTGGACGTTTACCGGAACCTCAATGTC 678
709 TACGCTCATGATTATCCAACAGGGGATGGCTCTGCGATCCGGGACTATATCCATGTGTATG 768
679 TATGACATGACTATCCCAACGAGATGGTAGTGGGTAAAGACTACATCAATGATGATG 738

FEATURES		source			
	1. .1313	/organism="Arabidopsis thaliana"			
		/mol_type="mRNA"			
		/db_xref="taxon:3702"			
		/chromosome="1"			
		/clone="RAFL09-23-D03 (R13909)"			
		/ecotype="Columbia"			
5'UTR	1. .1104				
CDS	105. .1160				
	/note="uridine diphosphate glucose epimerase"				
	/codon_start=1				
	/product="At1g12780/F13K23_21"				
	/protein_id="AAL06868.1"				
	/db_xref="GI:15809880"				
	/translation="MGSSVEQNILVTGGAGFICTHTVVQLLKQGFVKVSIIDPNSVIT EAYDRELVPGLSKLDFNLGRNKGDIKLFQKQFEDAVIHFAGLKAVGESVEN PRYFDNNILVTGINLAYETWAKYCKMVFSSATVYVQGPKEIPCMEDFELKAMPYGR TKLFLSEIARDIOKASPEWRKIILLRFNVPVGAHSGSIGEDPKGIENNLMPIYQQVAV GRPELVNVGHDYIGKGSNAVROYIHVMDLADGHI AALRLKLPADPKIGCTTAVNLGHTGO GTSLVSNVAAPFAFASGKKPIPIKLPRRSGDATVAYASTEKAKEELGHWKAKYGVDEMCR DQWKNANNPWGFGNKL"				
3'UTR	1161. .1313				
ORIGIN					
	Query Match	49.5%	Score 649.2;	DB 8;	Length 1313;
	Best Local Similarity	76.3%	Pred. No. 8.9e-147;		
	Matches 798;	Conservative 0;	Mismatches 248;	Indels 0;	Gaps 0;
QY	49	CAACACATTTCTGTCACCGGTGGTGGCGGTTTCATTGGCACCCACACCGCTGCTTACGCTT	108		
DB	123	CAGAACATTTCTTGTTACTGTGGTGTGCTGGGCTTTATCGGAGCGCATACTGTTGTTCAACTT	182		
QY	109	CTCAAGAGCTGGCTTCAGCGTTTCAATAATCGACAAATTTTCGATAACTCCGTCATGGAGCA	168		
DB	183	CTCAAGATGGTTTTAAGTTTTGATCATCGATATTTTGTATTAATCTGTATTCGAGACT	242		
QY	169	GTGACCGCGCTCCGCCAAGTGGTTGGCCCTCTGCTTTCTCAGAACCTTCCAAATTCACCCAG	228		
DB	243	GTTGATAGAGTTAGGAGGCTTGTGTGCTCGATCTCTCCAAGAAGCTCGACTTCAATCTG	302		
QY	229	GGCGATCTCGGAATAGGGATGACTTGGAGAACTCTTCTCCAAACAAACATTGTATGCC	288		
DB	303	GGTGATCTAAGAACCAAAAGGGGACATTGAGAAACTATTTCTCCAAGCAGAGATTGTATGCT	362		
QY	289	GTGATCCACATTTGTGCTCTTGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGCTCGCTAT	348		
DB	363	GTGATTCATTTTGGCGTCTTAAAGCTGTGGTGAGAGTGTGTGAACCCCTCGCGGTAC	422		
QY	349	TTTGATTTTAAATTTGGTTGGACCAATCAACCTCTTAAGTTTATGGCAAAAGTATATATGC	408		
DB	423	TTTGACAATAACTTGGTTGGAAACAATCAATCTATATGAGACCATGGCCAAAGTACAACTGC	482		
QY	409	AAAAAGATGTTTTTCTCATCATCTGCAACCGTTTTATGCGCAACCTGAAAAAGATACCGTGT	468		
DB	483	AAAATGATGTTGTTTTCAITCTTTCGCACCTGTTTATGCAACCACTGAAAAGATTCCAATGC	542		
QY	469	GAGGAGGATTTCAAGTTTCAAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTTGGAA	528		
DB	543	ATGGAAGACTTTGAAATTTAAAGGCTATGAAATCCCTATGCTGCTAAGCTCTTCTTGAA	602		
QY	529	GAAATTCGCCAGNATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGATAC	588		
DB	603	GAAATAGCTAGAGATATTCAAAAGGCGAAGACCGGAATTTGGAAGATTAATTCGCTGAGGTAC	662		
QY	589	TTCAATCCAGTTGGGGCTCATGAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCCCCA	648		

Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. .1420

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="11"

/clone="RAFL05-11-C04 (R16095)"

/ecotype="Columbia"

/note="This clone is in pBluescript"

gene

1. .1420

/gene="At1g12780"

/note="synonym: F13K23.3"

211. 1266

/gene="At1g12780"

/codon_start=1

/product="uridine diphosphate glucose epimerase"

/protein_id="AA053267.1"

/db_xref="GI:21539429"

/translation="MGSSVEONILVTGGAGFTGTHVQVLLKDGPKVSIIDNFNSVIV
EAVDRVRELVGLDLSKLDLNFGLRNKDIKLFKORFADVIHFAGLKAVGSVEN
PRYFDNNLVGINLYETWAKYCNKMPSSATYVGGPEKIPCKMDEFLKAMPEYGR
TKLLEIARDIOKAEPEWRIILIRFNPVGAHSGSTGEDPKGIPNNLMPIYQQVAV
GRLELVNVTGHDYPTEDGSAVRDIHVMDLADGHIALLRKLPADPKIGCTIYNLTGG
GTSLELVAAAFKASGKIPILKCPRRSGDATAVYASTKAELKELGWKANKYGVDEMCR
DQKWANNPNWPGYQNKI"

ORIGIN

49.5%; Score 649.2; DB 8; Length 1420;

Query Match 76.3%; Pred. No. 9e-147; Indels 0; Gaps 0;

Matches 799; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

49 CAACACATTCGTGTCACCGTGTGCGGTTTCATTGGCACCACACACCGTGTTCAGCTT 108
229 CAGAACATTCCTTACTGTTGTTGCTGCTGTTTATCGGACGCATATCTGTTTCAACT 288
109 CTCAAAGCTGGTTCAGCTTCAATTAATCGACATTTTCGTAATCTCCGTCTATGAGCA 168
289 CTCAAAGATGTTTAAAGTTTCGATCATGATAATTTGATACTCTGTTATCGAAGCT 348
169 GTGACCGCTCGCCAAAGTGTGTCCTCTGCTTCTCAGAACCTCCAAATTCACCCAG 228
349 GTTATAGATTAGGAGCTTGTGTTCTGATCTCTCCAGAGCTCGACTTCAATCTG 408
229 GCGATCTCCGAAATAGGATGACCTTGGAGAACTCTTCTCCAAACACATTTGATGCC 288
409 GGTGATCTAAGAAACAAAGGGGACATTTGAGAACTATCTCCAGCAGAGATTTGATGCT 468
289 GTGATCCACTTCTGCTGTTGAAAGCGTTGCTGAAAGGTTGCGAAGCCCGTGTCTAT 348
469 GTGATTCATTTGGGGTCTTAAAGCTGTGGGTGAGAGTTTGAAACCTCGCCGCTAC 528
349 TTTGATTTTAAATTTGGTGGCACCACCACTCTACAGTTTATGGCAAGTATATTTGC 408
529 TTTGACAAATTAATTTGGTTGGAACCAATCAATCTATATGAGACCATGTCGAAAGTAC 588
409 AAAAGATGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGT 468
589 AAAATGATGTTGTTTCTCATCTTCTGACCTGTTTATGGCAACCTGAAAGATTCATGC 548
469 GAGGAGATTTCAAGTTTCAAGTATGAATCCCTATGGACGGACCAAGCTTTTCCTGAA 528
649 ATGGAAGATTTGAAATTAAGGCTATGAATCTTATGTTGCTGTAAGCTCTTCTTGA 708
529 GAAATTCCTGAGTATTCAGAGCTGACACCAAGTATGAGATCATATTTACTGAGATAC 588
709 GAAATCTAGAGATATTCAAAAGGCGAACCAGGATGAGAAATTTCTGCTGAGGTAC 768
589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCA 648

769 TTCAATCTCTAGGACACATGAGCTGGCAGTATTGGTGGAGATCCAAAGGCATCCCC 828
649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGAGTTGACTGAACCTCAATGTA 708
829 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGAGTTGACTGAACCTCAATGTA 888
709 TACGGTCATGATTATCAACGAGGATGCTCTGCGATCGGAGCTATATCCATGTTGATG 768
889 TATGGACATGACTATCCACCGGAGTGTAGTGGCTTAAGAGCTACATCCATGTTGATG 948
769 GACTTGGCAGATGCGCATATTGCTGCGCTCGCAAGCTTTTCAACACGGAGAACATAGGT 828
949 GATTAGCAGATGCGCATATTGCTGCGCTCGCAAGCTTTTCTGATCCAAAGATTGGT 1008
829 TGTACTGTTTACAACTGGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
1009 TGTACTGCTTACAACTTAGGAGCTGTCAAGGAACGCTCTGTGTGTAGAAATGGTTGAGCT 1068
889 TTTGMAAGGCTTCTGCAAGAAATTTCCAGTAAATTTATGTCACAAAGACCGGGAGAT 948
1069 TTTGMAAGGCTTCTGCAAGAAATTTCCAGTAAATTTATGTCACAAAGACCGGGAGAT 1128
949 GCGACTGAGTTTATGCTATCTCAGAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAAC 1008
1129 GCAACAGCTTTTATGCTTCAACAGAGAGAGCTGAGAAAGAACTTGGCTGGAAGGCAAA 1188
1009 TATGGTGTGAGGAGATGTCAGGAGCAATGGAATTTGGCAAGAACTTCCCTGGGGT 1068
1189 TATGGTGTGAGGAGATGTCAGGAGCAATGGAATTTGGCAAGAACTTCCCTGGGGT 1248
1069 TACCGGGGAGGCTTGAATTAGCTT 1094
1249 TACCAGATAAGCTTTGAAATTTACTT 1274
AX507662 1056 bp DNA linear PAT 27-SEP-2002
Sequence 2357 from Patent WO0216655.
AX507662
AX507662.1 GI:23388899
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1
Harper, J.P., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 2357 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
Location/Qualifiers
1. .1056
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Query Match 49.1%; Score 644.4; DB 6; Length 1056;
Best Local Similarity 76.3%; Pred. No. 1.3e-145;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;

49 CAACACATTCGTGTCACCGTGTGCGGTTTCATTGGCACCACACACCGTGTTCAGCTT 108
19 CAGAACATTCCTTACTGTTGTTGCTGCTGCTTATCGGAGCATATCTGTTTCAACTT 78
109 CTCAAAGCTGGCTTCAAGTATGAGATTTGATGATGATGATGATGATGATGATGATGAT 168
79 CTCAAAGATGTTTAAAGTTTCGATCATCGATGATGATGATGATGATGATGATGATGAT 138

```
QY 169 GTGACCGCGTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCAATTCACCCAG 228
D 139 GTTGATAGAGTTAGGAGCTGTGTGCTGCTGATCTCTCCAGAAGCTGACATCAATCTG 198
QY 229 GCGGATCTCCGGAATAGGATGACCTGGAGAACTCTCTCCAAACACATTTTCATGCC 288
D 199 GGTGATCTAAGAAACAAAGGAGCATTTGAGAACTATTTCTCCAGCAGAGATTTGATGCT 258
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGGTGTCTGAAAGCGTTGCGAAGCCCGTCGCTAT 348
D 259 GTGATTCATTTTGGCGGTCTTAAAGCTGTGGGTGAGAGTGTGAAACCTCGCGCTAC 318
QY 349 TTTGATTTTAAATTTGGTGGCAACCTCAACCTCTACAGTCTTATGCGAAGTATTAATTC 408
D 319 TTTGCAAAATCACTTTGTTGGAACTCAATCTATATGAGACCATGCGAAGTACAACTGTC 378
QY 409 AAAAAAGATGTTTTCTTCATCATCTGCAACCGTTTTATGGCCCAACCTGAAAGATACCGTGT 468
D 379 AAAATGATGGTGTTCATCTTCCTGCACTGTTTATGGACAACCTGAAAGATTCATGTC 438
QY 469 GAGGAGATTTCAAGTTCAAGTATGAATCCCTATGAGCGGACCAAGCTTTTCCTGAA 528
D 439 ATGGAAGACTTTGAATTTAAGGCTATGAATCTCTTATGTCGTAAGCTCTTTCTTGA 498
QY 529 GAAATTTGCCGAGATATTCAGAAAGCTGAACAGCAATGGAAGATCATATTTACTGAGATAC 588
D 499 GAAATAGCTAGATATTTCAAAGGCAGAACCGGATGAGATTTATTTCTGCTGAGGTAC 558
QY 589 TTCAATCCAGTTGGGCTCATGAAATGGCAAACTCGGTGAAGATCCCAAGGGCATCCCA 648
D 559 TTCAATCCTGTAGGACACATGAGTGGCAGTATTTGTTGAGGATCCAAAGGCATCCCC 618
QY 649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATTTGACTGAACCTCAATGTA 708
D 619 AATAACTCATGCTTACATTCACCAAGTGGCGGTGAGCGTTTACCGGAATCAATGTC 678
QY 709 TACGGTCAATGATPATCAACGAGGATGCTCTGCGATCCGGGACTATATCCATGTGATG 768
D 679 TATGGACATGACTATCCCAACGAGGATGCTAGTGGCGTTGAGCGTTTACCGGAATCAATGTC 738
QY 769 GACTTGGCAGATGCGCATTTGTCGCCCTCGCAAGCTCTTCCACAGCGAGAACATAGT 828
D 739 GATTTAGCAGATGCGCATATGCTGCGCTCAGGAAGCTATTTGCTGATCCAAAGATTTGT 798
QY 829 TGTAAGCTTACAACTGGGAACTGCTCGTGGAACTCTGCTGTTGAAATGGTTACAGCA 888
D 799 TGTAAGCTTACAACTAGGACTGCTCAAGGAAGCTCTGTTAGAAATGTTTCAAGCT 858
QY 889 TTTGAAAGGCTTCTGCGAAGAAATTCAGTAAATTTATGTCAGAGAGACCGGAGAT 948
D 859 TTTGAAAGGCTTCTGCGAAGAAATTCCTGATTTAAGCTCTGTCGAGAGAGGTCAGGAGAT 918
QY 949 GCGACTGAGGTTTTATGATCTACAGAGAGAGCTGAGAAAGAACTGTTGGGAAGCAAC 1008
D 919 GCAACAGCAGTTTATGCTTCAACAGAGAGGCTGAGAAAGAACTTGGCTGGGAAGCAAA 978
QY 1009 TATGTTGTGAGAGATGTGACGGGACCAATGGAATTTGGGCAAGAACATCCCTGGGGT 1068
D 979 TATGAGATGATGAGATGTGACAGATCAGTGGAAATGGGCAACAAATTAATCCATGGGCT 1038
QY 1069 TACCGGGGAGGCTTGA 1086
D 1039 TACCAGAAATAGCTTTGA 1056
```

```
RESULT 7
AF334724
LOCUS AF334724 1056 bp mRNA linear PLN 07-MAY-2003
DEFINITION Arabidopsis thaliana clone C00128 (e) putative uridine diphosphate
ACCESSION AF334724
VERSION AF334724.1 GI:12248020
```

```
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
rosids; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1056)
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L.,
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R.,
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A.,
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
TITLE 2 (bases 1 to 1056)
JOURNAL Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
REFERENCE Goldsmith,A.D., Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q.,
AUTHORS Chen,H., Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and
Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2001) Plant Gene Expression Center, 800 Buchanan
REFERENCE Street, Albany, CA 94710, USA
AUTHORS 3 (bases 1 to 1056)
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R.,
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A.,
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan
COMMENT Street, Albany, CA 94710, USA
Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.
FEATURES
source
1..1056
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="C00128 (e)"
/ecotype="Columbia"
/notes="This clone is now in pUNI 51 vector. Previously,
it was in pUNI-T3-D/V5-His-TOPO under the clone number
C00128."
gene
1..1056
/gene="At-g12780"
CDS
1..1056
/gene="Atg12780"
/codon_start=1
/evidence=experimental
/product="putative uridine diphosphate glucose epimerase"
/protein_id="AA050102.1"
/db_xref="GI:12248021"
/translation="MGSSVEQNILVTGGAGFITHVTVQLLKDFKVSIIIDNPDNSVI
EAYDRVELVGLDLSKLDLNFKNKDIKLFKQRFDAVIHFAGLKAVGESVEN
PRYFDNLVGTINLVETMAKYNKMWVFSSATVYQPEKIFCMEDFELKAMPYGR
TKLFLEIARDIOKABPEWRIILIRYFNVPVAGSGSIGDEPKIPNNLMPYIQQAV
GRPELVNVDYDPTGDSAVRDYIHVMDLADGHIAALRKLFPADPKIGCTAYNLGTQ
GTSVLENVAFKASCKKIPKLPKPRSGDATAVYASTEKAKELGWKAKYGVDMCR
DQWLENNPNPWGTONKL"
Query Match 49.1%; Score 644.4; DB 8; Length 1056;
Best Local Similarity 76.3%; Pred. No. 1.3e-145;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
QY 49 CAACACATTTGTCACCGTGGTGGCGGTTTCAATGGCACCCACCGTCTGTCAGCTT 108
D 19 CAGAACATTTCTGTACTGTGGTGGCTTTATCGGAGCGCATACTGTGTTCACATTT 78
QY 109 CTCAAAGCTGCTTCAGCGTTTCAATATCGACAATTTTCGATAACTCCGTCATCGAAGCA 168
D 79 CTCAAAGATGTTTAAAGTTTCATCATCGATAATTTTGATAACTCTGTATCGAAGCT 138
```

QY 169 GTGGACCGCGTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACTCAATTCACCCAG 228
 Db 139 GTTGATAGAGTTAGGAGAGTTGTTGGTCTGATCTCTCCAAAGCTGGACTTCAATCTG 198
 QY 229 GCGGATCTCGGAATAGGATGACTTGGAGAACTCTCTCCAAACCAACATTTGATGCC 288
 Db 199 GGTGATCTAAGAAACAAGGGACATTTGAAACTATTTCTCCAGACAGATTGATGCT 258
 QY 289 GTGATCCATTTGCTGGCTTGAAGCGGTGTCTGAAGGTTTGCAGAGCCCGCTCGCTAT 348
 Db 259 GTGATTCATTTTGGCGGTCTTAAAGCTGTGGGTGAGAGTTTGAAGAACCTCGCGCTAC 318
 QY 349 TTGATTTTAAATTTGGTTCGACCATCAACCTCTAGAGTTTATGCGAAAGTATAATGC 408
 Db 319 TTGACAATTAATTTGGTTCGACCATCAATCTATATGACCATGCGAAAGTACAATGC 378
 QY 409 AAAAGATGGTTTTCATCTATCTCAACCGTTTATGCGCAACCTGAAAGATACCGTGT 468
 Db 379 AAAATGATGGTTTTCATCTCTGCACTGTTTATGGAACAACCTGAAAGATTCATGC 438
 QY 469 GAGGAGGATTTCAAGTTTCAAGCTATGATCTTATGCGACGACCAAGCTTTCTCTGAA 528
 Db 439 ATGGAAGATTTGAATTAAGGCTATGAATCTTATGCTGCTGACTTCTTTCTGAA 498
 QY 529 GAAATTCGCGAGATTTCAAGAACTGAAACGAGATGGAAGATCATATTTACTGAGATAC 588
 Db 499 GAAATAGCTAGATATTTCAAGAGCGAGAACCGGAATGGAGATTTATTTCTGCTGAGGTAC 558
 QY 589 TTCATCCAGTTGGGCTCATGAAAGTGCACAACTCGGTGAGATCCCAAGGCAATCCCA 648
 Db 559 TTCATCCGTGTAGGAGCATGAGAGTGGCAATTTGTTGAGATCCCAAGGCAATCCCC 618
 QY 649 AATAACCTCATGCTTTACATTTCAAGAACTGAGTGTGTTGGAAGATTTGACTGAATCAATGTA 708
 Db 619 AATAACCTCATGCTTTACATTTCAAGAACTGAGTGTGTTGGAAGATTTGACTGAATCAATGTC 678
 QY 709 TACGCTCATGATTTCCACAGAGGATGCTCTGCGATCCGGGACTATATCATGTGATG 768
 Db 679 TATGGACATGACTATCCCAAGGATGTTGCTGCGGTAAAGAGACTACATCCATGTGATG 738
 QY 769 GACTTGGCAGATGGCCATATTTGCTGCTCGCAAGAGCTTTCCACAAAGAGCAATAGGT 828
 Db 739 GATTTAGCAGATGGCCATATTTGCTGCTCAGAGAGCTATTTGCTGATCCAAAGATTGT 798
 QY 829 TGTACTGCTTAACTGGAAGTCTGCTGTGGAACATCTGCTGCTTGAATGTTACAGCA 888
 Db 799 TGTACTGCTTAACTGGAAGTCTGCTGGAAGAGCTGCTGTTAGAAATGTTGCACT 858
 QY 889 TTTGAAAGGCTTCTGGCAAGAAATTTCCAGTAAATTTATGTCGAAGAGACCGGGAGAT 948
 Db 859 TTTGAAAGGCTTCTGGCAAGAAATTTCCGATTAAGCTCTGTCGAGAGAGTCCAGAGAT 918
 QY 949 GCGACTGAGTTTATGATCTTACAGAGAGCTGGAAGAACTTTGTTGGAAGGCAAAAC 1008
 Db 919 GCAACAGCAGTTTATGCTTCAACAGAGAGGCTGGAAGAACTTTGCTGGAAGGCAAAA 978
 QY 1009 TATGTTGTCGAGAGATGTCAGGAGCAATCGAATTCGGCAAGAAACAATCCCTGGGT 1068
 Db 979 TATGAGTGTGAGATGTCAGAGATCAGTGGAAATGGCAACAATAATCCATGGGT 1038
 QY 1069 TACGCGGGAAGCTTGA 1086
 Db 1039 TACCAGAAATAGCTTTGA 1056

RESULT 8

ATUDGE
 LOCUS ATUDGE 1356 bp mRNA linear PLN 11-FEB-1998
 DEFINITION A.thaliana mRNA for uridine diphosphate glucose epimerase.
 ACCESSION Z54214
 VERSION Z54214.1 GI:1143391
 KEYWORDS uridine diphosphate glucose epimerase.
 SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 REFERENCE 1 (bases 1 to 1356)
 AUTHORS Newman,T., de Bruijn,F.J., Green,P., Keegstra,K., Kende,H.,
 McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,
 Retzel,E. and Somerville,C.
 TITLE Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106 (4), 1241-1255 (1994)
 JOURNAL 95148729
 MEDLINE 7846151
 REFERENCE 2 (bases 1 to 1356)
 AUTHORS Dormann,P. and Bemmig,C.
 TITLE Functional expression of uridine 5'-diphospho-glucose 4-epimerase
 (EC 5.1.3.2) from Arabidopsis thaliana in Saccharomyces cerevisiae
 and Escherichia coli
 Arch. Biochem. Biophys. 327 (1), 27-34 (1996)
 JOURNAL 96201343
 MEDLINE 8615692
 REFERENCE 3 (bases 1 to 1356)
 AUTHORS Doermann,P.P.
 TITLE Direct Submission
 JOURNAL Submitted (20-SEP-1995) Peter Doermann, Institut fuer
 Genbiologische Forschung Berlin, GmbH, Ihnestrasse 63, Berlin,
 Berlin, 14195, Germany
 COMMENT Overlapping sequence: T22968 (sequence obtained from same clone).
 FEATURES
 source
 1..1356
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /variety="Col-2"
 /db_xref="taxon:3702"
 /clone="10862077, GenBank ID T22968"
 /tissue_type="mixture of etiolated seedlings, roots,
 rosette leaves, stems, flowers, siliques"
 /clone_lib="PRL-2 in lambda Ziplox, ABRC Ohio State,
 Columbus, OH 43210, USA"
 66
 /citation=[1]
 /replaces="n"
 151..1206
 /EC_number="5.1.3.2"
 /codon_start=1
 /evidence="experimental"
 /product="uridine diphosphate glucose epimerase"
 /protein_id="CAA90941.1"
 /db_xref="GI:1143392"
 /db_xref="GOA:Q42605"
 /translation="MGSSVEQNILVTGGAGFTGTHVTVOLLKDGPKVSIIDPDSVI
 EAVDRVRELVGPDLSKKLDFNLGDKNKGDIKLFQKDFDAVHFAGLKAVGESVEK
 GRFYFNNTLVGTINLYETWAKYCKMVFSSATVYQPEKIPCMEDFLKAMNPTGR
 TKLFLSEIARDIKAEPEWRIILRYFNVPVGHESGIGEDPKGIPNNILMPTVQVAV
 GRLEPNVYGHDPYPTEDGSAVRDYTHVMDLADHIAALKLPADPKIGCTAVNLGTGQ
 GTSVLRWAAFEKSKGKTIPIKLCPRRSGDATAVAVASTEKAELKELGKAKYGVDENCR
 DQWKAFNPNWYQNKI"
 217
 /citation=[1]
 /replaces="n"
 299
 /citation=[1]
 /replaces="g"
 312..313
 /citation=[1]
 /replaces="g"
 319..320
 /citation=[1]
 /replaces="n"
 329..330
 /citation=[1]
 /replaces="g"
 conflict
 conflict
 conflict
 conflict
 conflict

```
conflict 338 /citation=[1]
          /replace=[1]
          /replace="nn"
conflict 345. 347 /citation=[1]
          /replace="atctgn"
conflict 353 354 /citation=[1]
          /replace="t"
conflict 357 /citation=[1]
          /replace="n"
conflict 360 /citation=[1]
          /replace="g"
conflict 371 /citation=[1]
          /replace="g"
conflict 373 374 /citation=[1]
          /replace="g"
conflict 384 /citation=[1]
          /replace="t"
conflict 398 399 /citation=[1]
          /replace="n"
conflict 403 /citation=[1]
          /replace="t"
conflict 410 /citation=[1]
          /replace="t"
conflict 412 413 /citation=[1]
          /replace="na"
conflict 416 417 /citation=[1]
          /replace="tt"
conflict 421 /citation=[1]
          /replace="c"
conflict 426 /citation=[1]
          /replace="g"
conflict 430 /citation=[1]
          /replace="t"
conflict 433. 447 /citation=[1]
          /replace="agccctg"
conflict 449 /citation=[1]
          /replace="g"
conflict 451 /citation=[1]
          /replace="n"
conflict 454 /citation=[1]
          /replace="g"
conflict 457. 458 /citation=[1]
          /replace="tt"
conflict 460. 467 /citation=[1]
          /replace="ttnaaacc"
conflict 469 /citation=[1]
          /replace="n"
conflict 471 /citation=[1]
          /replace="g"
conflict 473
```

```
          /citation=[1]
          /replace="c"
          /citation=[1]
          /replace="gn"
          /citation=[1]
          /replace="nt"
          /citation=[1]
          /replace="mncan"
          /citation=[1]
          /replace="nt"

ORIGIN
```

```
Query Match 49.0%; Score 642.8; DB 8; Length 1356;
Best Local Similarity 75.9%; Pred. No. 3.2e-145;
Matches 794; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

QY 49 CAACACATTCGTCCACGGTGGTCCGGTTTCATTGGSCACCCACACCGTCGTTCACTT 108
Db 169 CAGAACATTCGTTACTGTGGTGGCTGTTATCGGAGCGCATCTGTGTCACTT 228
QY 109 CTCAAAGCTGCTTCAGCGTTTCAATAATCGACAATTCGATACTCCGTCATCGAAGCA 168
Db 229 CTCAAAGATGTTTTAAGGTTTCGATCATCGATAATTTGATAACTCTGTATCGAAGCT 288
QY 169 GTGACCGCTCCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAG 228
Db 289 GTTGATAGAGTAGGAGCTGTGTGGTCCCTGATCTCTCAAGAAAGCTCGACTCAATCTG 348
QY 229 GCGATCTCCGGAATAGGATGACTTGGAGAAATCTTCTCCAAACAAACATTTGATGCC 288
Db 349 GGTGATCTAGAAACAAGGGACATTTGAGAACTATTTCTCCAGCAGAGATTTGATGCT 408
QY 289 GTGATCCACTTGTGCTGTTGAAAGCGTTGCTGAAAGCGTTGCGAAGCCCGCTCGCTAT 348
Db 409 GTGATTCATTTTGGGGTCTTAAAGCTGTGGTGAGAGTGTGTAAGAGGGTTCGCGCTAC 468
QY 349 TTTGATTTTAAATTTGGTGGCACCATCAACTCTACGAGTTTATGGCAAGTATAATTGC 408
Db 469 TTTGACATTAATTTGGTGGTGGACAACTATATATGAGCATTGGCAAGTACACTGC 528
QY 409 AAAAAGATGTTTTCTCATCATCTGCAACCGTTTTATGGCAACCTGAAAAGATACCGTGT 468
Db 529 AAAATGATGTTTTCTCATCTTCTGCCACTGTTTTATGGACAACCTGAAAAGATTCATGC 588
QY 469 GAGGAGGATTTCAAGTTACAGCTATGAATCCCTATGAGCGGACCAAGCTTTTCTCGAA 528
Db 589 ATGGAAGACTTTGAATTAAGGCTATGAACTTATGATCCCTATGTTGTTAGTACTTTCTTGA 648
QY 529 GAAATTCGCCGAGATATTTCAGAAAGCTGAACCAAGATGGAAGATCATATTACTGAGATAC 588
Db 649 GAAATAGCTAGAGATATTCAAAGGCGAAGACCGGAATGGAGAAATTATCTGCTGAGGTAC 708
QY 589 TTCATCCAGTTGGGCTCATGAAGTGGCAACTCGTGAAGTCCCAAGGGGATCCCA 648
Db 709 TTCATCTGTAGGAGCAGATGAGAGTGGCAGTATTGGTGAAGGATCAAAGGATATCCCC 768
QY 649 AATAACCTCATGCTTACATTTCAGCAAGTAGCTGTTGGAAGATTCAGTGAATCAATGTA 708
Db 769 AATAACCTCATGCTTACATCCAAAGTGGCGCTTGGAGCTTTACCGGAATCAATGTC 828
QY 709 TACGCTCATGATTATCAACGAGGATGGCTCTCGATCCGGGACTATATCCATGTATG 768
Db 829 TATGACATGACTATCCCAAGGATGGTAGTCCGTAAGAGACTACATCATCATGTATG 888
QY 769 GACTTGGCAGATGGCCATATTGCTGCCCTCGAAAGCTTCTTCAACCGGAGAAATAGGT 828
Db 889 GATTTAGCAGATGGCCATATCGCTGCGCTCAGGAAGCTATTTGCTGATCCCAAGATTTGT 948
QY 829 TGTACTGCTTACAAACCTGGGAACTGGTGGTGGAAACATCTGTGCTTGAATGGTTACAGCA 888
```

```

Db 949 TGTACTGCTTCAATCTAGGACCTGGTCAAGGAACGCTGTGTAGAAAATGGTTGCAGCT 1008
Qy 889 TTTGAAAAGGCTTCTGGCAAGAAATCCAGTAAATATATCTCCAAAGAACCGGGAGAT 948
Db 1009 TTTGAAAAGCTTCCGGCAAGAAATCCGATTAAGCTCTCTCCGAGAGTCAAGGAT 1068
Qy 949 GGGACTGAGGTTATGCACTCTACAGAGAGACTGAGAAAGAACTGGTTGGAAGGCAAC 1008
Db 1069 GCAACAGCAGTTTATGCTTCAACAGAGAGAGCTGAGAAAGAACTTGGCTGGAAGCAAAA 1128
Qy 1009 TATGCTGTGGAGGAGATGTGCAAGGACCAATGGAAATGGGCAAGAAACAATCCCTGGGGT 1068
Db 1129 TATGAGTGGATGAGATGTGAGAGATCAGTGGAAATGGGCTTTCAATAATCCATGGGT 1188
Qy 1059 TACGCGGGAGCCTTGAATAGCTT 1094
Db 1189 TACCAGAATAAGCTTTGAATTACTT 1214

RESULT 9
AK117913
LOCUS
DEFINITION
Arabidopsis.thaliana At1g63180 mRNA for putative uridine
diphosphate glucose epimerase, complete cds, clone: RAF119-09-L10.
ACCESSION
AK117913
VERSION
AK117913.1 GI:26450887
KEYWORDS
FLI CDNA; CAP trapper.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y., and Shinozaki, K.
Arabidopsis thaliana full-length cDNA
2 (bases 1 to 1263)
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y., and Shinozaki, K.
Direct Submission
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail: msekigsc.riken.go.jp,
URL: http://pfweb.gsc.riken.go.jp, Tel: 81-45-503-9625,
Fax: 81-45-503-9586)
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified paluescript vector.
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for
further details.
FEATURES
Location/Qualifiers
1..1263
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF119-09-L10"
/note="common name: thale cress"
1..1263
/gene="At1g63180/F16M19_8"
52..1107
/codon_start=1
/product="putative uridine diphosphate glucose epimerase"
/protein_id="BAC42551.1"

```

```

/db_xref="GI:26450888"
/translation="MGSSVEQNILVTGGAGFIGHTVTVVQLLNQGRKVTIIDLNSVY
EAVRYRELVDGLSTKLEFNLGNLDLRNKDIEKLFNSQRFDAIVHFAGKAVGESVGN
PRYFDNLVGTINLYETMAKYNCKOMYFSSATVYQGFVPCVPCVPELOQMAYVGR
TKLFLIEARDTHAAEPWKIILLFYFVGAHESGRIGEDPKGI:PNLMPYIQVAV
GRLPLNVGHDYPTMDGSVRDYIHNMDLADGHVAALNKLKLFSDSKIGTANLTGQ
GTSVLEMYSSPFKASGKKIPIKLCPRRAGDATAVAVASTQKAEKELGWAKYGVDSMCR
DOWNWANKNPWFQKPK"
ORIGIN
Query Match 48.1%; Score 631; DB 8; Length 1263;
Best Local Similarity 75.2%; Pred. No. 2.3e-142;
Matches 787; Conservative 0; Mismatches 260; Indels 0; Gaps 0;
Qy 49 CAACACATCTCGTCACCGCTGTCGCGCTTTTCATTGGCACCCACACCGCTCGTTCAGCTT 108
Db 70 CAGAACATTTTGTGACTGTGTGCTGGATTCATTGGNACACATACCTGTTGTTAGCTT 129
Qy 109 CTCAAAGCTGGCTTCAGCGTTTCAATAATCGACAATTTGATAAATTCGATCGAAGCA 168
Db 130 TTGAATCAGGGTTTAAAGGTTACGATCATTTGATAAATCTCTGCTCGTTGAAGCT 189
Qy 169 GTGGACCGGCTCGCCACGCTGTTGGCCTCTGCTTCTCAGAAGCTCCAAATCACCCAG 228
Db 190 GTTCATAGGGTTAGGGAATCTTGTGCTGCTGATCTCTTACCAGCTTGAATCAATCTG 249
Qy 229 GCGCATCTCCGGAATAGGAGATGACTTGGAGAAATCTTCTCCAAAACAACATTTGATGCC 288
Db 250 GGTGATCTAAGAAACAAGAGAGATATTGAGAAATCTTTTCCAAATCAGAGATTTGATGCT 309
Qy 289 GTGATCAGTTTGTGCTTGAAGGGTGTCTGAAAGGTTGCGAAGGTTGCGAAGCCCGCTGCTAT 348
Db 310 GTGATTCATTTGCTGCTCTTAAAGCTGGGAGAAAGTGTGGAAACCTCTGCTGCTTAC 369
Qy 349 TTTGATTTTAAATTTGGTGGCCACCTCAACCTCTACGAGTTTATGGCAAGATATATTCG 408
Db 370 TTTGATAATACTTAGTTGGAATCATCATCTATGAGACCATGSCAAAATACAACTGC 429
Qy 409 AAAAAGATGGTTTCTCATCTGCAACCGTTTATGGCAACCTGAAAGATACCGTGT 468
Db 430 AAAATGATGGTGTGTTGCTGCTGCTGCAACAGTTTATGGTCAACCTGAAATAGTCCCATGT 489
Qy 469 GAGGAGATTTCAAGTTTACAAGCTATGAATCCTATGAGCGGACCAAGCTTTTCTGTGAA 528
Db 490 GTGGAAGACTTTGAGTTACAGCTATGAATCTTATGGTCTGCTACTAAGCTGTTCTTGAA 549
Qy 529 GAAATGCCGAGATATTGAGAAAGCTGAAACAGAGATGGAAGATATATATCTGAGATAC 588
Db 550 GAGATAGCTAGACACATTCACGCTGCGGAACAGAGATGGAAGATATATCTTCTGAGGTAC 609
Qy 589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGGAAGATCCCAAGGGCATCCCA 648
Db 610 TTTCAATCCGGTTGAGCTCAGAGATGGAAGATTTGGAAGATCCCAAGGGCATACCG 669
Qy 649 AATAAAGCTCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATTTGACTGAATCATATGA 708
Db 670 AATAATCTCATGCTTATATCCAAAGTGGGGTTGGAAGATTTACTGAACTCAATATGA 729
Qy 709 TACGCTCATGATTTATCCACAGGGGATGCTCTGGATCCGGGACTATATCCATGTGATG 768
Db 730 TTTGGACATGATTTATCTTACCAGTGTAGCGGGTTAGAGACTATCATCATGTAATG 789
Qy 769 GACTTGGCAGATGGCCATATTTGTCCTTCGCAAGCTCTTCCACACGGAGAACATAGGT 828
Db 790 GATTTAGCAGATGGCCATGTGGCTGGTTAAACAAATTTGTTTCAGACTCAAAGATTGGC 849
Qy 829 TGTACTGCTTACAACTGGGAATGCTGCTGGAGACATCTGCTTGAATTTGGTTACAGCA 888
Db 850 TGTACTGCTTATAATCTTGCACTGCTCAAGGAATCTTGTCTTGAATTTGGTTCTTCT 909
Qy 889 TTTGAAAAGCTTCTCGCAAGAAAATTCAGTAAATATATGTCCAAAGAGACGGGAGAT 948
Db 910 TTTGAAAAGCTTCTCGCAAGAAAATACCTATCAAGCTATGTCCAAAGAGAGCTGAGAT 969

```

```

QY 949 GCGACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTGGAGGCAAC 1008
Db 970 GCAACAGCTGTTTATGCTTCACTCAGAAAGCTGAGAAAGAACTTGGTGGAGGCAAA 1029
QY 1009 TATGGTGTGAGGAGATGTGACGGGACCAATGGAATGGGCAAGAAACAAATCCCTGGGGT 1068
Db 1030 TATGGAGTTGATGAGATGTGACAGATCAATGGAATGGGCAAGAAATGAATCCATGGGGT 1089
QY 1069 TAGCGGGGAGCTTGAATACCTTG 1095
Db 1090 TTCAGAAGAAGCCTTGATGATTG 1116

RESULT 10
AY085887
LOCUS AY085887 1397 bp mRNA linear PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 1932 mRNA, complete sequence.
ACCESSION AY085887
VERSION AY085887.1 GI:21404597
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1397)
AUTHORS Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE 22088475
PUBMED 12093376
REFERENCE 2 (bases 1 to 1397)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1397)
AUTHORS Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE Direct Submission
JOURNAL Submitted (11-WAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genet carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.
FEATURES
Location/Qualifiers
1..1397
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="1932"
197..1252
/codon_start=1
/product="uridine diphosphate glucose epimerase, putative"
/protein_id="PAM63099.1"
/db_xref="GI:21554018"
/translation="MGSSVEQNILVTGGAGFIGTHTVQLLNQGGFKVITIDLDNSV"

```

ORIGIN

Query Match 48.0%; Score 629.4; DB 8; Length 1397;

Best Local Similarity 75.1%; Pred. No. 5.8e-142; Indels 0; Gaps 0;
Matches 786; Conservative 0; Mismatches 261;

```

QY 49 CAACACATTTGCTCACCGGTGGTCCGGTTTCATTTGGCACCCACACCGTGTTCAGCTT 108
Db 215 CAGAATATTTTGGTGACTGGTGGTCTGGAATTCATTGGAAACACATACCTGTGTTCAGCTT 274
QY 109 CTCAAAGCTGGCTTCAGCGTTTCATATCGACAAATTCGATTAATTCCTCGTCATCGAGCA 168
Db 275 TTGAATCAGGGTTTTTAAGGTTACGATCATGTATATCTTGTATAACTCTGTGCTTGAAGCT 334
QY 169 GTGACCGCGTCCGCGCAAGTGGTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAG 228
Db 335 GTTCATAGGGTTAGGGAACCTTGTGGTCTCTGATCTCTCTACCAAGCTTGAATTCATCTG 394
QY 229 GCGCATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGCC 288
Db 395 GGTGATCTTAAGAAACAAGGAGATATTGAGAACTCTTTTCCAAATCAGAGATTTGATGCT 454
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGTGTCTGAAAGCGTTGGAAAGCGTGGAAACCTCGTGT 348
Db 455 GTGATTCACTTTGTGCTTAAAGCTGTGGAGAAAGTGTGGAAACCTCGTGTGTAC 514
QY 349 TTTGATTTAATTTGGTGGCACCATCAACTCTACGAGTTTATGGCAAGTATTAATTCG 408
Db 515 TTTGATAATAACTTAGTTGGAACTATCAATCTATATGAGACCATCGGCAAAATACAACTGC 574
QY 409 AAAAGAGATGTTTTCTCATCATCTGCAACCGTTTTATGCCAACCTGAAAGATACCGGT 468
Db 575 AAAATGATGGTGTGTTTGTCTGCAACAGTTTATGCTCAACTGAAATAGTCCCATGT 634
QY 469 GAGGAGATTTCAAGTTTACAGCTATGAATCCCTATGAGCGGACCAAGCTTTCTCTGGAA 528
Db 635 GTGGAAGACTTTGAGTTACAGGCTATGAATCCTTATGCTGCTACTAAGCTGTGTTCTTGA 694
QY 529 GAAATGCCAGATATTTCAGAAAGCTGAAACAGAAATGGAAGATCATATTACTCAGATAC 588
Db 695 GAGATAGTAGAGACATTCACGCTGCGAACCAGAAATGGAAGATTAATCTCTGAGGTAC 754
QY 589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCCGTGGAAGATCCCAAGGGCATCCCA 648
Db 755 TTCAATCCGGTTGGAGCTCACGAGATGGAAGAAATTCGAGAAGATCCAAAGGGCATACCG 814
QY 649 AATAACCTCATGCTTACATTCACCAAGTAGCTGCTTGGAGATTTGACTGAACCTCAATGTA 708
Db 815 AATAATCTCATGCTTATATCCAAACAGTGGCGGTGGAAGATTACCTGAACCTCAATGTA 874
QY 709 TACGCTCATGATTATCCAAACGAGGATGGCTCTCCGATCCGGGACTATATCCATGTGATG 768
Db 875 TTTGGACATGATTATCCTACCATGATGGTAGCGGGTTAGAGACTACATCCATGTAATG 934
QY 769 GACTTGGCAGATGGCCATATTGCTGCCCTGCGAAAGCTCTTCACACGGAGAACATAGGT 828
Db 935 GATTTAGCAGATGGCCATGTGGCTGCGGTAAACAAATTTGTTTTCAGACTCAAGATTTGGC 994
QY 829 TGTAATGCTTACAACTCCGGGAACCTGGTCTGTGGAACATCTGTGCTTGAATGTTCACAGCA 888
Db 995 TGTAATGCTTATATCTTGGCACTGGTTCAGGAACCTCTGTCTTAGAATGTGTTTCTCT 1054
QY 889 TTTGAAAGGCTTCTGGCAAGAAAATTCAGTAAATTAATGTCCAGAGAGACCGGAGAT 948
Db 1055 TTTGAAAGGCTTCTGGCAGAAAATATACCTATCAAGCTATGTCCAGAGAGAGCTGGAGAT 1114
QY 949 GCGACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTGGAGGCAAC 1008

```


QY 1008 CTATGTTGGAGGAGATGTCAGGACCAATGGAAATGGGCAAGAAACAATCCCTGGG 1067
 Db 998 ATATGAGTTGATGAGATGTCAGAGATCAATGGAATGGGCAATAAAGATCCATGGG 1057
 QY 1068 TTACGGCGGAGACCTTGAATAGCTTG 1095
 Db 1058 TTTCAGAGAGACCTTGATGATTG 1085

RESULT 12

AK104977

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:001-124-A06, full

insert sequence.

AK104977.1 GI:32990186

VERSION

FLI CDNA; oligo capping.

Oryza sativa (japonica cultivar-group)

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1

The Rice Full-Length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team;

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

Kojima, K., Shishiki, T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group; Otono, Y., Tsunoda, Y.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, N., Xie, Q., Lu, M.,

Nariikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikeda, R., Ishibiki, J., Kawana, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,

Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,

Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

12869764

2 (bases 1 to 1375)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,

Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Itoh, M., Kagawa, I.,

Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kawai, J.,

Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,

Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,

Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,

Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K.,

Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,

Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,

Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,

Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,

Tagawa, A., Takahashi, E., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,

Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and

Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of

Agrobiological Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,

Tel:81-29-838-7007, Fax:81-29-838-7007)

TITLE

JOURNAL

REFERENCE

AUTHORS

AUTHORS

AUTHORS

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdnaol.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,

Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,

Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and

Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,

Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,

Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,

Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center

and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,

Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,

Ota, Y., Saitoh, H., Sakai, C., Sakazume, N., Sano, H.,

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,

Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,

Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1375

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clones="001-124-A06"

ORIGIN

Query Match 36.8%; Score 482.8; DB 8; Length 1375;

Best Local Similarity 66.8%; Pred. No. 2.4e-106;

Matches 688; Conservative 0; Mismatches 342; Indels 0; Gaps 0;

QY 54 CATTCTGTGTACCGGTGGTGGCGGTTTCATTGGCACCACACACCGTCTCAGTTCTCAA 113

DB 122 CGTGTGTGTACCGGTGGTGGCGGTTTCATTGGCACCACACACCGTCTCAGTTCTCAA 181

QY 114 AGCTGGCTTCAGCGTTTCATATTCGACATTCGATACTCCGTCATCGGAGCAGTGA 173

DB 182 GAAGGGGTTCGCGGTTCACCGTCTCGTCAAACTCCCAACTCCCGCGGAGGCGCTCGA 241

QY 174 CGCGTTCGCGCAAGTGGTGGCGGCTCTCTCTCAGAACTCCCAACTCCCGCGGAGGCGCTCGA 233

DB 242 CGCGTTCGCGCTCATCGCGGCGCGGCTCTCTCTCAGAACTCCCAACTCCCGCGGAGGCGCTCGA 301

QY 234 TCTCCGGAATAGGATGATCGTTGGAGAACTCTCTCTCCAAACACATTTGATCGGTGAT 293

DB 302 TCTCAAGACGAAGGACGATCGGAGAGGTGTTCGCCGCCAAGAGGTATGACGCGGTAT 361

QY 294 CCACCTTCGCTGGCTTGAAGCGGTGTCTGAAAGCGGTTCGGAAGCGGCTCGCTATTTTGA 353

DB 362 CCACCTTCGCGGCTTGAAGCGGTGTCTGAAAGCGGTTCGGAAGCGGCTCGCTATTTTGA 421

QY 354 TTTTAATTTGGTGGCAACCACTCTCTACAGTTTATGCGCAAGTATATTTGCAAAA 413

DB 422 GAACAACGTCGCGGCAACCATGAACTCTCTCTCGGCCATGACCAAGTACGGCTCGCAAA 481

QY 414 GATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGA 473

DB 482 GATAGTGTCTCGTCTCGGCGAGCGTGTAGCGGCGGCGGAGAGACCCCTCGTCTG 541

QY 474 GGATTTCAGTTTACAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTGGAGAAAT 533

DB 542 GGATTTCAGTTTACAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTGGAGAAAT 601


```
QY 54 CATTCTGCTACCGGTGTCGGGTTTCAATTCGGACCCACACCGTCTGCTTCAAGTCTCTCAA 113
Db 178 CGTCTGCTGAGCGGGGAGCGGGGTACATCGGACGCGACACCGTCTGCTCTGCTCGA 237
QY 114 AGCTGCTTTCAGGCTTCAATTAATGACAAATTCGTAATTCGTAATTCGTAATTCGTA 173
Db 238 GAAGGGTTCGGGTACCGTCTGTCGACAACTTCACAACTTCGTCGCGGAGCGCTCGA 297
QY 174 CCGGTCGCCCAAGTGTGTCGCTCTGCTTCTCAGAACCTCCAAATTCACCCAGGGCGA 233
Db 298 CCGGTCGCCCTCATCGCGGGCGCGCTCTCCGCGCGCTTCGATTCATCGCGGGCGA 357
QY 234 TCTCCGGAATAGGATGCTCGAGAACTCTCTCCAAACAACTTGTATGTCGCTGAT 293
Db 358 TCTCAAGACAGGACGATGAGHAGTGTTCGCGCCAGAGATGACCGCGTAT 417
QY 294 CCACTTTGCTGCTGGAACCGGTTCTCGAAAGCGTTGCGAACCCCGCTGCTATTTTGA 353
Db 418 CCACTTCGCGGCTGAAGCGGTGGGAGAGCGTTCGCGCACCCCGCAGATGATCTACGA 477
QY 354 TTTTAATTTGGTTGGCACCATCACTCTACAGTTTATGCGCAAGTATATTTGCAAAA 413
Db 478 GAACAACGTCGCGGCGCACCATGAACTCTACCGCATGACCAAGTACGGCTGCAAGAA 537
QY 414 GATGTTTTCTCATCATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGCTGTGAGGA 473
Db 538 GATAGTGTCTCGTCTCGCGACGCTGTA CGGCCAGCGGAGAGACCCCTCGCTCGA 597
QY 474 GATTTCAAGTTCAAGCTATGATCCCTATGAGCGGACCAAGCTTTCTCGGAAGAAT 533
Db 598 GGATTTCAAGCTGAGCGCTCTCAACCCATACGCGCACCAAGAGCTCGCTCGGAGACTA 657
QY 534 TGCCCGAGATATTCAGAAAGCTGAACAGATGGAAGATCATATTACTGAGATATCTCAA 593
Db 658 CTTCGCGAGGTGACGCGCGCCGACCGCGAGATGAGGTGATCTCTGCTCAGTACTTCA 717
QY 594 TCAGTTGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGAGATCCCAAAATA 653
Db 718 CCCATCGCGCTCACCGGAGCGGACATCGGGAGGAGCCCGAGGGGATCCCAACAA 777
QY 654 CTTATGCTTACATTCAGCAAGTACGTTCTGGAAGATGACTGAACTCAATGATATAGG 713
Db 778 CTTCTTCGTAATCCAGAGGTGCGCTCGCGCCCGCCCGAGCTCAAGCTCTACGTTACG 837
QY 714 TCATGATTTCAACAGGAGTGGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
Db 838 CGTCGACTACCAACAGGAGCGCACCGGATCAGGATTTACATATGATGAGCT 897
QY 774 GGCAGATGGCCATATTGCTGCGCTCGAAAGCTCTTCAACAGGAGACATAGTTGTAC 833
Db 898 TGCCGATGGCCCATTTGCGGCACTGGAGAGCTCTTCGCTACTCTCTGACATTTGTTGT 957
QY 834 TGCTTACAACTGGGAACCTGGTCTGGAACATCTGTGCTTGAATGGTTTACAGCATTTGA 893
Db 958 GGTTCATCTAGGAACAGGTTGTGGAACACCGGTCTCGAGTGGTGAAGCGTTTGA 1017
QY 894 AAGGCTTCTGCAAGAAATTCAGTAAATTTATGTCGAAGAGCCGGGAGATGCGAC 953
Db 1018 GAGGGGTTCGGAAGAAATTCCTATCAAGATTTGCCCCAGAGACCTGGAGATTGAC 1077
QY 954 TGAGGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGTGGAAGGCAAACTATGG 1013
Db 1078 TGAGGTTTACCTTCACTGACAGGCGCAGAGAGAGCTCGATGAGATGCTCGGTTGG 1137
QY 1014 TGTGAGGAGATGTCAGGAGCAATGGAATTTGGCAAGAACTTCCCTGGGTTACGC 1073
Db 1138 AATAGAGGACATGTCAGGAGCCAGTGGAAATTTGGGCAAGAGAAATCCCTACGATACAG 1197
QY 1074 GCGGAGCCT 1083
Db 1198 CGCCAATGCT 1207
```

```
RESULT 14
AB097460
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) OsUGE-3 mRNA for putative
UDP-glucose 4-epimerase, complete cds.
ACCESSION
AB097460
VERSION
AB097460.1 GI:26106013
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 Suzuki, K. and Kitamura, S.
Cloning of UDP-glucose 4-epimerase genes in Oryza sativa
Unpublished
2 (bases 1 to 1223)
Suzuki, K. and Kitamura, S.
Direct Submission
Submitted (03-DEC-2002) Kiyoshi Suzuki, Osaka Prefecture
University, Graduate School of Agriculture and Biological Sciences;
Gakuen-cho, 1-1, Sakai, Osaka 599-8531, Japan
(B-mail:ksuzuki@biochem.osakafu-u.ac.jp,
URL:http://www.biochem.osakafu-u.ac.jp/BPC/MAIN-J.html,
Tel:81-72-254-9457, Fax:81-72-254-9458)
FEATURES
Location/Qualifiers
1..1223
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/tissue_type="immature seed"
/country="Japan: Kyoto, Shimogamo"
1..1223
/gene="OsUGE-3"
65..1147
/gene="OsUGE-3"
/codon_start=1
/product="putative UDP-glucose 4-epimerase"
/db_xref="GI:26106014"
/db_xref="GI:26106014"
/translation="MYGNGEGRKAGASVLVTGGAGVIGTHVLRLEKGFVTVVDN
FHSVPEALDRVLIAGALSARLDPIAGLKSDDMEKVFPAKRYDAVHIFAGLAV
FNSVHPQYNNVAVGKIKIYFSSATGATQPKPTCPVEDSKLSA
LNFYPTKLVLENYFRQVADPEMVIILRRYDPIGAHSGDIBSDPRGIPNNLIPY
IQVAVRRPELVYGVDPTRDGTARDYIHVVDLADGHAALKEFLPATPDTCVAY
NLGTGCTTVLEVVKAPEEASGKKIPIKICPRPBGDTVEVYASTDKAKELGHSAREG
IEDCRDQNWAKNPFVGSANAEQN"
ORIGIN
Query Match 36.6%; Score 479.6; DB 8; Length 1223;
Best Local Similarity 56.6%; Pred. No. 1.4e-105;
Matches 686; Conservative 0; Mismatches 344; Indels 0; Gaps 0;
QY 54 CATTCTGCTACCGGTGCTCGGTTTCAATTCGGACCCACACCGTCTGCTTCAAGTCTCTCAA 113
Db 106 CGTCTGCTGACGGGGAGCGGGGTACATCGGACGCGACACCGTCTGCGGCTGCTGGA 165
QY 114 AGCTGCTTTCAGCGTTTCAATAATTCGCAANTTCGATAACTCCGTCATGGAAGCAGTGA 173
Db 166 GAAGGGTTCGCGGTCAACCGTCGTCGACAACTTCCAACTCCGTCGCCGAGGCGCTCGA 225
QY 174 CCGGTCGCCCAAGTGTGTCGCTCTGCTTCTCAGAACCTCCAAATTCACCCAGGGCGA 233
Db 226 CCGGTCGCCCTCATTCGCGGCGCGCGCTCTCGCCCGCTCGACTTCTATCGCCGGGA 285
QY 234 TCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACAACTTGTATGTCGCTGAT 293
Db 286 TCTCAAGACAGGACGACATGGAAGAGTGTTCGCGCCAGAGAGTATGACGCGCTGAT 345
QY 294 CCATTTCGCTTGAAGCGTTCGTAAGAGGTTGCTGAAGGTTGCGAGCCCGTCTGCTATTTGA 353
```

Db 346 CCACCTTCGCGGCTCAAGCGCTGGGGAGAGCGTCGCGCACCGCAGATGTACTACA 405
 Qy 354 TTTTAAATTTGGTGGACCAATCAACTCTACGAGTTTATGSCAAAGTATAATTCAGAAA 413
 Db 406 GAACAACTGTCGGGACCACTGCTCTACCTCCGCAATGACCAAGTACGCTGCAAGAA 465
 Qy 414 GATGTTTCTCATCTATGCAACCGTTTATGSCCAACTGAAAGATACCGTGTGAGGA 473
 Db 466 GATAGTGTCTCTGTCGCGACGCGCTACGCGCAGCGGAGAACCCCTCTGCTCGA 525
 Qy 474 GATTTTCAAGTTTACAAGCTATCAATCCCTATGAGCGGACCAAGCTTTCTTGGAGAAAT 533
 Db 526 GATTTCCAACTGAGCGCTCTCAACCCATACGCGACCAAGCTCTGCTCTGGAGAACTA 585
 Qy 534 TCCCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATPATCTGAGATCTCAA 593
 Db 586 CTTCCGGCAGGTGCAGCGCGCCGACCCGAGATGAGGCTGATCTCTGCTCAGTACTTCA 645
 Qy 594 TCCAGTTGGGCTCATGAAGTGGCAACTCGGTGAAGATCCCAAGGCGATCCCAATAA 653
 Db 646 CCCCATCGCGCTCAACGAGCGGCGACATCGGGAGGACCCCGGGGATCCCAATAA 705
 Qy 654 CTTCTATGCTTACATTCAGCAAGTATGTTTGGAAAGATTGACTGAACCTCAATGTATACGG 713
 Db 706 CTTCTCTCGTACATCCAGCAGTCTGCGCTGCGCGCCGCGCGGCTCAAGCTCTACGG 765
 Qy 714 TCATGATTAACAAGAGGATGGCTCTGCGATCCGGGACTATATCTCATGTGATGGACTT 773
 Db 766 CBTGACTACTACCAACAGGACCGGACCGCGATCAGGGATATACATACATGATGAGCCT 825
 Qy 774 GGCAGATGCGCATATTTGCTGCGCTCGAAAGCTCTTTCACACGAGAGACATAGGTTGTAC 833
 Db 826 TGCGATGCGCAATTTGCGCACTGGAGAGCTCTTCGCTACTCTGCTGACATGGTTGTGT 885
 Qy 834 TGCTTACAACTGGGAACTGGTCTGGGAACTCTGCTGTTGAAATGGTTACAGCATTTGA 893
 Db 886 GGCCTTACAATCTAGGAACAGGCTGTGGAACAACCGTCTCGAGGTGGTGAAGCGGTTCA 945
 Qy 894 AAGAGCTTCTGCAAGAAATTTCCAGTAAATTTATGTCACAGAACCGGAGATGCGAC 953
 Db 946 GNGGGCTCGGAAAGAAATTTCTATCAGATTTGCCCGACAGACCTGGAGATTGAC 1005
 Qy 954 TGAGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAAACTATGG 1013
 Db 1006 TGAGTTTATCGCTTCCACTGCAAGGCCAAGAGGAGCTCGGATGGAGTGTCTCGGTTGG 1065
 Qy 1014 TGTGGAGAGATGTGCGAGGACCAATGGAATGGGCAAGACATCCCTGGGTTACGC 1073
 Db 1066 AATAGAGGACATGTGCGAGGACCAAGTGAATTTGGCCAAAGAAATCCGTACGATACAG 1125
 Qy 1074 GGGAGAGCCT 1083
 Db 1126 CGCCATGCT 1135

RESULT 15
 BT009044
 LOCUS
 DEFINITION
 Triticum aestivum clone wdk5c.pk006.04:fls, full insert mRNA
 1516 bp mRNA linear PLN 20-JUN-2003
 sequence.
 ACCESSION
 BT009044
 VERSION
 BT009044.1 GI:32128595
 KEYWORDS
 FLI-CDNA.
 SOURCE
 Triticum aestivum (bread wheat)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1516)
 Tingley, S.V., Wolters, P., Powell, W., Dolan, M., Miao, G.-H.,
 Caraher, N.R., Hanafey, M.K. and Hainey, C.F.
 Direct Submission
 Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and

Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
 USA
 FEATURES
 source
 Location/Qualifiers
 1..1516
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wdk5c.pk006.04:fls"
 ORIGIN
 Query Match 36.3%; Score 476.4; DB 8; Length 1516;
 Best Local Similarity 66.4%; Pred. No. 8.8e-105;
 Matches 684; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
 Qy 54 CATTCGGTTCACCGGTGGTCCCGTTTCATTTGGCACCCACACCGTCTGATCGTCTTCAA 113
 Db 12 CGTCTGGTTCACCGCGCGCGGGTTCATTCGGCACGCAACCGTCTGATCGTCTGGA 71
 Qy 114 AGCTGGCTTCAGCGTTTCAATATACCAATTTTCGATAATCTCGTTCATGGAACGATGGA 173
 Db 72 GAAGGGCTACCGCTCACCGCGTTCGCAACTTCGCAACTCGTCCCGAGCGCTCGA 131
 Qy 174 CCGCGTCCGCAAGTGGTTGGCCCTCTGCTTCTCAGAACTCTCAAACTCACTCACTCACTCA 233
 Db 132 CCGCGTCCGCAATCGTGGCGCCCGCTCTCGCGCGCTCTCAATTCATCTTCGCGGA 191
 Qy 234 TCTCCGGAATAGGATGATCTTGAGAACTCTTCTCCAAACAACTTTGATGCGCGTAT 293
 Db 192 CTTGACGATCAAGATGATCTTGAGAAAGTCTTTCGCGCGCAAGATGATGACGCGCTGAT 251
 Qy 294 CCATTTTGTGGTTCGAAAGCGGTCTGAAAGCGTTCGAAAGCGTTCGAAAGCGTTCGTTGA 353
 Db 252 ACATTTCCGCGGCTCAAGCGGTGGCGGAGAGCGTGGCGCACCCCGAGATGTACAACCG 311
 Qy 354 TTTTAATTTGGTGGCACCACTCAACCTCTACGAGTTTATGGCAAGTATAATTTGCAAAA 413
 Db 312 CAACAACATCGTCCGACCGCTCAACCTCTACGAGCTCATGAAGAGCAAGGTTGCAACAA 371
 Qy 414 GATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGA 473
 Db 372 GTTGGTGTCTCTGCTCGCGACCGTGTACGCGCAGCGGAGAGTGGCTGCTTCA 431
 Qy 474 GATTTCAAGTTACAGCTATGAACTCCCTATGAGCGGACCAAGCTTTCTCTGAGAAAT 533
 Db 432 GGAATCCCGCTCAAGCGCTCAACCGTACCGGAGGACCAAGCTGTACCTGGAGGAT 491
 Qy 534 TGCCCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATTTACTGAGATACTTCAA 593
 Db 492 GCTGCGGACTACAGCAACCGGAGTGGAGGACGATCTCTGCTGCTACTTCAA 551
 Qy 594 TCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCAATCCCAATAA 653
 Db 552 CCGCATCGCGCACACGAGAGCGCGACATCGGGGAGGACCCCAAGGGCGTCCCAACAA 611
 Qy 654 CCTCATGCTTTACATTCAGCAAGTAGCTGTGGAAGATTGACTGAATCTCAATGTATACGG 713
 Db 612 CTTGCTCCCTCATCATCAGCAGGTGGCGCTCGCGCCCGCGCCCGAGCTCAACCTTACGG 671
 Qy 714 TCATGATTAATCCCAAGAGGATGCTCTGATTCGCGGACTATATTCATGTGTATGAGCTT 773
 Db 672 CCAGACTACCGCACCCGCGACCGGACCCGCGTCAAGGAGTCAATCCATCGTGTGACCT 731
 Qy 774 GGCAGATGGCCATATTTGCTGCGCTCGAAAGCTTTTCAACAGGAGAACATAGGTTGTAC 833
 Db 732 CGCGAGCGGCACATCGCGCGCTCGAGAGCTTTTCGCCACCCCTGACATCGGCTGTGT 791
 Qy 834 TGCTTACAACTGGAACTGGTCTGGTGAACATCTGTCTTGAATGGTTACAGATTTGA 893
 Db 792 GGCATACAACTGGGACCGGGCGGGACGAGCTGTGGAGTGTGGAGTGTGAGCGCTTCA 851
 Qy 894 AAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCAGAGAGACCGGAGATCGGAC 953
 Db 852 GAAGGCAATACGGCAAGAAATTCGCGTGAAGATGTGCCCGAGGCGCGGCGATTCGA 911

Qy	954	TGAGGTTTATGCACTACAGAGAGAGCTGAGAAAGAACTTGGTTGGAGGGCAAACTATGG	1013
Db	912	GCAGGTGTACGCGTCCACCGCCCAAGGCCGAGAGAGGAGCTCGGCTGGAGGGCCCAAGTACGG	971
Qy	1014	TGTGGAGGAGATGTGACAGGACCAATGGAATTGGGCAAGAAACAATCCCTGGGGTTACGC	1073
Db	972	GATCGAGGAGATGTGACAGGACCAATGGAATTGGGCAAGAAACAATCCCTGGGGTTACGC	1031
Qy	1074	GGGAAAGCCT	1083
Db	1032	CGGCAACGCT	1041

Search completed: March 22, 2004, 19:14:59
Job time : 5330 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:28:33 ; Search time 572 Seconds
(without alignment)
9744.119 Million cell updates/sec

Title: US-09-913-064a-13
Perfect score: 1312
Sequence: 1 gcacagcactctctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%
Maximum Watch 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1312	100.0	1312	3	AA59439 Nucleotid
2	644.4	49.1	1056	6	AB214552 Arabidops
3	629.4	48.0	1397	3	AAC34854 Arabidops
4	603.2	46.0	1488	6	ABA00335 Potato ps
5	601.8	45.9	1056	6	ABA00336 Potato ps
6	544.2	41.5	1508	3	AAC37107 Arabidops
7	476.4	36.3	1516	3	AA59440 Nucleotid
8	445	33.9	1065	7	ADA69621 Rice Gene
9	445	33.9	1498	3	AA59442 Nucleotid
10	431.4	32.9	1532	3	AA59443 Nucleotid
11	431.2	32.9	1394	3	AAC51240 Arabidops
12	431.2	32.9	1396	3	AAC39471 Arabidops
13	428.4	32.7	1286	3	AAC36878 Arabidops
14	424.8	32.4	479	3	AA59435 Nucleotid
15	423.6	32.3	1292	3	AAC47345 Arabidops
16	416	31.7	1314	3	AAC50699 Arabidops
17	412	31.4	1314	3	AAC47111 Arabidops
18	393.4	30.0	1393	3	AA59441 Nucleotid
19	385.6	29.4	1335	7	ADA70308 Rice Gene
20	312	23.8	641	6	ABQ66038 Arabidops
21	278.2	21.2	4626	2	AA584321 Stealth v
22	277.8	21.2	1017	6	ABN99338 B. subtil
23	277.8	21.2	5793	6	AA517998 Plasmid p

24	277.8	21.2	5943	6	AA517997
25	267.6	20.4	1023	7	ACF70498
26	267.6	20.4	69727	7	ACF65374
27	267.6	20.4	110000	7	ACF67367_35
28	262.4	20.0	1008	7	ABX07603
29	262.4	20.0	110000	7	ABX07603 S. pneumo
30	262.2	20.0	1050	8	ADB09547
31	262.2	20.0	110000	8	ADB12064_09
32	261.8	20.0	100848	4	AAF28552_09
33	261.6	19.9	1017	7	ABX07359
34	261.6	19.9	13206	2	AAV52166
35	261.6	19.9	110000	7	ABX56454_13
36	261.6	19.9	110000	7	ABX56454_14
37	248.6	18.9	545	3	AA57605
38	247.2	18.8	3140	2	AAV52291
39	244	18.6	1047	3	AAZ45319
40	244	18.6	1488	6	ABL42872
41	244	18.6	1488	6	ABL67513
42	244	18.6	1488	6	ABL63090
43	244	18.6	1488	6	ABL67063
44	241.4	18.4	110000	2	AAT42063_03
45	240.8	18.4	1249	2	AAT58301

ALIGNMENTS

RESULT 1
AAA59439
ID AAA59439 standard; DNA; 1312 BP.

XX AC AAA59439;

XX 14-NOV-2000 (first entry)

XX DE Nucleotide sequence of UDP-galactose 4-epimerase.

XX KW UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level; ss.

XX OS Glycine max.

XX FH Key Location/Qualifiers
XX CDS 415..1086

XX FT /*tag= a
XX FT /product= "UDP-galactose 4-epimerase"
XX FT /note= "the codons encoding amino acids 3-129 are not given"

XX WO200047755-A2.

XX PD 17-AUG-2000.

XX PF 09-FEB-2000; 2000WO-US003453.

XX PR 10-FEB-1999; 99US-0119588P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Carlson TU, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;

XX PI Rafalski JA, Thorpe CU;

XX DR WPI; 2000-549152/50.

XX DR P-PSDB; AAB07819.

XX PT Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant cells to e.g. decrease cell wall constituents and increase starch content.

XX PS Claim 2; Page 51; 61pp; English.

XX CC The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is

CC responsible for the interconversion of UDP-glucose and UDP-galactose. As
 CC the activity of the enzyme appears to be particularly limiting to carbon
 CC flux into the raffinose/saccharide pathway, reduction of its activity should
 CC decrease the levels of raffinose and stachyose in seeds. The
 CC polynucleotide is useful to transform plant cells to enhance the degree
 CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
 CC the expression of UDP-galactose-4-epimerase within a cell enables
 CC modification of plant carbohydrate metabolism. This allows growth of
 CC grains with reduced cell wall constituents (fiber) and increased levels
 CC of starch
 XX

SQ Sequence 1312 BP; 386 A; 273 C; 300 G; 353 T; 0 U; 0 Other;

Query Match 100.0%; Score 1312; DB 3; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGCCACTTCTCTCCCTCTCTATTGCGACATGGTGTCTCTCCCAACACATTTCTG 60
 DB 1 GCACGAGCCACTTCTCTCCCTCTCTATTGCGACATGGTGTCTCTCCCAACACATTTCTG 60
 QY 61 GTCACGGTGGTGGCGGTTTCAATGCGACCCACACACCGTCTGATCTCTCAAGCTGGC 120
 DB 61 GTCACGGTGGTGGCGGTTTCAATGCGACCCACACACCGTCTGATCTCTCAAGCTGGC 120
 QY 121 TTCAGCGTTTCAATATGACAAATTCGATACTCCGTCATGGAGCAGTGGACCGGTC 180
 DB 121 TTCAGCGTTTCAATATGACAAATTCGATACTCCGTCATGGAGCAGTGGACCGGTC 180
 QY 181 CGCCAAAGTGGTGGCGGTTTCTCTCAGAACCTTCAATTCACCCGGGCGATCTCCGG 240
 DB 181 CGCCAAAGTGGTGGCGGTTTCTCTCAGAACCTTCAATTCACCCGGGCGATCTCCGG 240
 QY 241 AATAGGATGACTTGGAGAACTCTCTCCCAAACTTCAATTTGATGATTTTAAAT 300
 DB 241 AATAGGATGACTTGGAGAACTCTCTCCCAAACTTCAATTTGATGATTTTAAAT 300
 QY 301 GCTGGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCCCGTGGCTATTTTGAATTTAAAT 360
 DB 301 GCTGGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCCCGTGGCTATTTTGAATTTAAAT 360
 QY 361 TTGGTTGGACCATCACTCTACAGTTTATGGCAAGATATATGCAAAAGATGGTT 420
 DB 361 TTGGTTGGACCATCACTCTACAGTTTATGGCAAGATATATGCAAAAGATGGTT 420
 QY 421 TTCTCATCTCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGATTC 480
 DB 421 TTCTCATCTCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGATTC 480
 QY 481 AAGTTACAAGCTATGAATCCCTATGGACGACCAAGCTTTCTGGAAGAAATTCGCCGA 540
 DB 481 AAGTTACAAGCTATGAATCCCTATGGACGACCAAGCTTTCTGGAAGAAATTCGCCGA 540
 QY 541 GATATTCAAGAGCTGAACAGAGATGGAAGATCATATTACTGAGATCTTCAATCCAGTT 600
 DB 541 GATATTCAAGAGCTGAACAGAGATGGAAGATCATATTACTGAGATCTTCAATCCAGTT 600
 QY 601 GGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAAAATTAACCTCATG 660
 DB 601 GGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAAAATTAACCTCATG 660
 QY 661 CCTTACATTCAGCAAGTAGCTCTGGAGATGATGTAAGTCAATGATATCGGTATGAT 720
 DB 661 CCTTACATTCAGCAAGTAGCTCTGGAGATGATGTAAGTCAATGATATCGGTATGAT 720
 QY 721 TATCAACGAGGATGGCTCTCGATCCGGGACTATATCCATGTGATGGACTTGGCAGAT 780
 DB 721 TATCAACGAGGATGGCTCTCGATCCGGGACTATATCCATGTGATGGACTTGGCAGAT 780
 QY 781 GGCATATTCTGCGCTCGGAAGCTTTCACAGCGGAGACATAGGTGACTGCTTAC 840
 DB 781 GGCATATTCTGCGCTCGGAAGCTTTCACAGCGGAGACATAGGTGACTGCTTAC 840

QY 841 AACCTGGGAACGTGTCGTGGAACATCTGTGCTTGAATGGTTTACAGCAATTTGAAAGGCT 900
 DB 841 AACCTGGGAACGTGTCGTGGAACATCTGTGCTTGAATGGTTTACAGCAATTTGAAAGGCT 900
 QY 901 TCTGGCAAGAAAATTCCTAGTAAATATGTCCCAAGACCGGAGATGCGACTGAGGTT 960
 DB 901 TCTGGCAAGAAAATTCCTAGTAAATATGTCCCAAGACCGGAGATGCGACTGAGGTT 960
 QY 961 TATGCAATCTACAGAGAGCTCGAGAAAGAACTTGTGTGGAAGGCAAACTATCGTGTGGAG 1020
 DB 961 TATGCAATCTACAGAGAGCTCGAGAAAGAACTTGTGTGGAAGGCAAACTATCGTGTGGAG 1020
 QY 1021 GAGATGTCAGGACCCCAATGGAATTTGGCAAGACCAATCCCTGGGTTTACCGGGGAG 1080
 DB 1021 GAGATGTCAGGACCCCAATGGAATTTGGCAAGACCAATCCCTGGGTTTACCGGGGAG 1080
 QY 1081 CCTTGAATTTAGCTTGAGAAATATCTGCTCATCTAGCAATGCTTTTCAATAAATAGGCA 1140
 DB 1081 CCTTGAATTTAGCTTGAGAAATATCTGCTCATCTAGCAATGCTTTTCAATAAATAGGCA 1140
 QY 1141 TCTCTTATATAGATATCTTTTATGTTGATGATTTTGGAGTCTGTTAGGAGTCTGTTGATAATCT 1200
 DB 1141 TCTCTTATATAGATATCTTTTATGTTGATGATTTTGGAGTCTGTTAGGAGTCTGTTGATAATCT 1200
 QY 1201 TGACAATAAAAAATTTGGCAGCATTTTCAAGAGTAAAGCTATGTATTTAAACAATAACTT 1260
 DB 1201 TGACAATAAAAAATTTGGCAGCATTTTCAAGAGTAAAGCTATGTATTTAAACAATAACTT 1260
 QY 1261 TAAATTAGACTGGCCATTGATTTGATTTGAAAAAATAAAAAAATAAAAAA 1312
 DB 1261 TAAATTAGACTGGCCATTGATTTGATTTGAAAAAATAAAAAAATAAAAAA 1312

RESULT 2

ABZ14552
 ID ABZ14552 standard; DNA; 1056 BP.

XX ABZ14552;

XX 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2357.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

PD 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and
 XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 2357; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant

CC cell has been exposed, comprising: (a) contacting nucleic acid

CC representative of expressed polynucleotides in the plant cell with an

CC array or probes representative of the plant cell genome; and (b)

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-014320P.
PR 12-JUL-1999; 99US-0143277P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160860P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

48.0%; Score 629.4; DB 3; Length 1397;

	Best Local Similarity	75.1%;	Pred. No. 1.1e-151;	
	Matches	786;	Conservative	0; Mismatches 261; Indels 0; Gaps 0;
QY	49	CAACACATTCCTGGTCA	CGGTGGTGGCGGTTTC	CAITGGCACCCACACACCGTCTCAGCTT 108
Db	215	CAGATAATTTTGGTGA	CTGGTGGTGGATT	CATTGGAAACATACATGTTGTCAGCTT 274
QY	109	CTCAAAAGCTGGTTC	AGGTTTCAATTAATCG	CAATTTGATTAATCTCGTCAATGAAGCA 168
Db	275	TGGAATCAGGGTTT	AAAGTTACGATCAIT	GAATTAATCTTGATAACTCTCTCGTGAAGCT 334
QY	169	GTGGACCGCTCG	CGCAAGTGGTGGCCCT	CTGCTTCTCAGAACTCTCAATTCACCCAG 228
Db	335	GTTCATAGGTT	TAGGAACTTGTGGT	CTCTGATCTCTCTACCAAGCTTCAATTCATCTG 394
QY	229	GCGCATCTCCGGA	ATAGGATGATCTGG	AGAAATCTTCTTCCAAAACAATTTGATGCC 288
Db	395	GGTGATCTAAGAA	CAAGAGAGATAT	TGAGAAATCTTTTCCAAATCAGAGATTTGATGCT 454
QY	289	GTGATCCACTTTG	CTGGCTTCAAGACGGT	TGCTGAAAGGTTGCGAAGCCCGTCTGCTAT 348
Db	455	GTGATTCATTTG	CTGGTCTTAAAGCT	GTGGGAAAGTGTGGAAACCCCTCGTGGTTAC 514
QY	349	TTTGATTTTAA	TTGGTTGGCAAC	CAACTCTACGAGTTTATGSCAAAGTATAATTCG 408
Db	515	TTTGATAATA	TAATCTAGTTGGA	ACTATCAATCTATAGACCAATGCAAAATACAATGCG 574
QY	409	AAAAAGATGGTTT	CTCATCATCTGGA	ACCGTTTATGGCCAACTTGAAAGATACCGTGT 468
Db	575	AAAAATGATGGT	TTTTGCTGCTGCA	ACAGTTTATGGTCAACCTGAATATGTCCTCATGT 634
QY	469	GAGGAGGATTTCA	AGCTATCAATCCCT	TATGGACGGAACCAAGCTTTTCTCTGGAA 528
Db	635	GTGGAAAGCTT	TGAGTTACAGCT	ATGAATCTTATGGTCTGATTAAGCTTTTCTTGA 694
QY	529	GAAATGGCCGAG	ATATTCAGAAAGCT	GAAACGAGATGGAAGATCATATTAATGAGATAC 588
Db	695	GAGATAGCTAG	AGACATTCAGCT	GCGGAAACCGAATGGAAGATATTTCTCTGAGGTAC 754
QY	589	TTCAATCCAGT	TGGGGCTCATGAA	AGTGGCAACTCGTGAAGATCCCAAGGGCATCCCA 648
Db	755	TTCAATCCGGT	TGGAGCTCAGAG	AGTGGAAAGATTCGAAGGCAATCGG 814
QY	649	AATAACCTCAT	GCCTTACATTCAG	CAAGTATGCTTTGGAGATGATGATCAATGTA 708
Db	815	AATAATCTCAT	GCCTTATATCCAA	CAAGTGGCGGTTGGAAGATTACCTGAACTCAATGTA 874
QY	709	TACGCTCATG	ATATCCAAAGAGG	ATGGCTCTGCGATCCGGACCTATATTCATCTGATG 768
Db	875	TTTGACATG	ATATCTCATCAT	GATGGATGGATGGCGGGTTAGACCTACATCTGAATG 934
QY	769	GACTTGGCAG	ATGGCCATATTC	GCTCGCCCTCGAAAGCTCTTCAACGAGAGAACATAGGT 828
Db	935	GATTTAGCAG	ATGGCCATATG	CGCTGTAAACAAATTTGTTTCAGACTCAAAAGATTGGC 994
QY	829	TGTACTGCTT	ACAACCTGGGA	ACTGGTGGTGGAAACATCTGTGCTTGAATATGTTACAGCA 888
Db	995	TGTACTGCTT	ATAATCTTTGG	CACTGGTCAAGGAACTCTGTCTTGAATATGTTCTTCT 1054
QY	889	TTTGGAAAG	CGCTTGGCAAGAA	ATTTCCAGTAAATATGTTCCAAAGAACCCGGAGAT 948
Db	1055	TTTGGAAAG	CTTGGCAGAA	ATACCTATCAAGCTATGTTCCAGAGAGCTCGAGAT 1114
QY	949	GCGACTGAG	GTATATGCATCT	ACAGAGAGCTGAGAAAGATCTGGTTGGAGGCAAC 1008
Db	1115	GCAACAG	CTGTTATGCTT	CAACTCAGAAAGCTGAGAAAGAACTTTGGCTGGAGGCAAAA 1174
QY	1009	TATGTGTGG	AGGAGATCTG	CAGGAGCAATGGAATTTGGCAAGAAACAATCCCTGGGCT 1068
Db	1175	TATGAGT	TGATGATGTG	CAGAGATCAATGGAACTGGSCAAATTAAGAAATCCATGGGGT 1234
QY	1069	TACGCGGG	GAAAGCCCTT	GAAATTAGCTTG 1095

Db 1235 TTCAGAGAGCCTTGATTGTTTG 1261

RESULT 4
 ABA00335
 ID ABA00335 standard; cDNA; 1488 BP.
 XX
 AC ABA00335;
 XX
 XX 10-DEC-2002 (first entry)
 XX
 DE Potato psen-1 gene.
 XX
 KW Gene; potato; psen-1; 5'-diphospho-galactose 4-epimerase; transgenic;
 XX uridine diphosphate glucose epimerase; UDPG; carbohydrate metabolism;
 XX Controlled Environment Agriculture; nutritional profile; DGE; ss.
 XX
 OS Solanum tuberosum.
 XX
 PH Key Location/Qualifiers
 FT CDS 143..1198
 FT /*tag= a
 FT /product= "DGE"
 XX
 XX WO200270649-A2.
 XX
 XX 12-SEP-2002.
 XX
 XX 01-FEB-2002; 2002WO-US002727.
 XX
 XX 01-FEB-2001; 2001US-0265311P.
 XX
 XX (DAIZ/) DAI Z.
 XX
 XX Dai Z, Shi L, Hooker BS;
 XX
 XX WPI; 2002-698747/75.
 XX
 XX P-PSDB; AAG79562.
 XX
 XX New isolated polynucleotide molecule comprising a gene coding for the
 PT enzyme 5'-diphospho galactose 4-epimerase, useful for regulating
 PT carbohydrate metabolism or nutritional profile in transgenic plants.
 XX
 PS Claim 1; Fig 1; 38pp; English.
 XX
 CC This sequence represents the full length cDNA sequence of potato psen-1
 CC which encodes the enzyme 5'-diphospho-galactose 4-epimerase (DGE). This
 CC sequence is homologous to the sequence of Arabidopsis thaliana uridine
 CC diphosphate glucose (UDPG) epimerase gene (86% homology at peptide level,
 CC 71% homology at the nucleotide level). The DGE polynucleotide molecules
 CC are useful in regulating carbohydrate metabolism in transgenic plants.
 CC The polynucleotide can be used in Controlled Environment Agriculture that
 CC employs an integrated system for commercial production of transgenic
 CC plants in a controlled environment. Controlling the level of DGE in
 CC transgenic plants is useful for regulating the nutritional profile of the
 CC plant
 XX
 SQ Sequence 1488 BP; 458 A; 240 C; 324 G; 466 T; 0 U; 0 Other;
 Query Match 46.0%; Score 603.2; DB 6; Length 1488;
 Best Local Similarity 73.3%; Pred. No. 6.4e-145;
 Matches 773; Conservative 0; Mismatches 283; Indels 0; Gaps 0;
 Qy 50 AACACATTCGTCACCGGTGGCGGTTTCATTGGCACCCACACCGTCGTTGAGCTTC 109
 Db 162 AAAATATTTTGGTTACTGGAGGAGCTGGTTTCATTGGACACACACTGTGGTGAGGTAC 221
 Qy 110 TCAAAGCTGGCTTCAGCGTTTCAATAATGCACAAATTCGATAATCCCGTCATGGAAGCAG 169
 Db 222 TGAATGAGGGTTCAAGTTACCATTCATTGATAATTCATTGTTGGAAGAGCTG 281
 Qy 170 TGGACCGCGTCGCCAAGTGGTTTGGCCCTCTGTTTCTCAGAACCTTCCAATTCACCACG 229


```
Db      617  TGGAGGATTTGAATTTGAAGGCTATGAATCTTATGGTGAACAAAGCTATTTCTTGAAG 558
Qy      530  AATATGCCGAGATATTCAGAAAGCTGACCAAGATGGAAGATCATATTTACTGAGATCT 589
Db      557  ATATTTCTCGGATATCCAGAGGCGAGATCAAGATGGAATATCATATTTGAGGTATT 498
Qy      590  TCAATCCAGTTGGGGTCTATGAAGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCAA 649
Db      497  TCAACCCAGTAGGAGCTCATGAAGTGGCAAACTCGGGAAGATCCAAAGGCGATCCCA 438
Qy      650  ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATGAGTCAATCATATGAT 709
Db      437  ACAATCTTATGCTTACATTCAGCAAGTAGCTGTTGGAAGATGAGTCAATCATATGAT 378
Qy      710  ACGGTCTATCATATCAACAGAGGATGGCTCTGCGATCCGGAATATATCCATGATGATGG 769
Db      377  ATGGCAACAGCTACCTTACCTGATGGTACCGGATAGAGATATATCCATGTTTTGG 318
Qy      770  ACTTGCAGATGGCCATATGTCCTGCGCTCGCAAAAGCTTTTCAACGGAGAACATAGGTT 829
Db      317  ATTAGCGGACGCTCATGTTGTTGCACTTCAGAGACTTCTAAGGCGAGATCATTTAGGTT 258
Qy      830  GTACTCTTACACCTCGGGAAGCTGTCGTTGGAACATCTGTCTTGAATGGTTACAGCAT 889
Db      257  GTGTTGCCATATTTGGGTACTGGAAGGCAAAATCTGTCTAGAGATGGTTGCTGTT 198
Qy      890  TTGAAAGAGCTTCTGCAAGAAATATCCAGTAAATATATGTCACAGAACCGGAGATG 949
Db      197  TTGAAAGAGCGTCTGAAAGAAATATCCGCTTAAATATGTTCAAGAGAACCGGAGATG 138
Qy      950  CGACTGAGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAAACT 1009
Db      137  CCAGTCTGTTTATGATCTACTGAAAGAGCTGAGAAAGAGCTCGGTGGAAGGCAAAAT 78
Qy      1010  ATGGTGTGAGGAGATGTGAGGAGCAATGGAATGGCAAGAACATCCCTCGGGTT 1069
Db      77  ATGGTATAATGAGATGTGAGGAGCACTGGAATGGCAAGAACATCCCTCGGGTT 18
Qy      1070  ACCGGGGAGCCTTGA 1086
Db      17  ACCAATCAAAGCCTTGA 1
RESULT 6
AAC37107
ID AAC37107 standard; DNA; 1508 BP.
XX AC AAC37107;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16200.
XX XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX XX
OS Arabidopsis thaliana.
XX PN EF1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX XX
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
```

```
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136352P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
```


Db 822 AATAACCTCATGCTTACATCCAAAGTGGCGTTGGAGCTTTACCGGAACCTCAATGTC 881
Qy 709 TAGGTCATGATATCCACGAGGATGCTCTGGATCCGGACTATATCCATGATG 768
Db 882 TAGGACATGACTATCCACCGAGGATGATGCGGTAAGAGACTTACATCCATGTGATG 941
Qy 769 GACTTGGCAGATGGCCATATGTCGCCCTCGGAAAGCTTTCCACACGAGAGAACATAGT 828
Db 942 GATTTAGCAGATGGCCATATGCTGGCTCAGGAGCTATTTGCTGATCCAAAGATTTGT 1001
Qy 829 TGTACTGCTTACACTGGGAACTGCTGTGGAACATCTGCTTGAATGTTTACACA 888
Db 1002 TGTACTGCTTACAACTAGGACTGCTCAAGGAACGCTGTGTGAATGTTGACGT 1061
Qy 889 TTTGAAAAGCTTCTGGCAAGAAAATCCAGTAAATATATGTCGAAGAGACCGGAGAT 948
Db 1062 TTTGAAAAGCTTCCGGCAAGAAAATCCGATTAAGCTCTGTCCGGAAGGTCAGGAGAT 1121
Qy 949 G 949
Db 1122 G 1122

RESULT 7
AA59440
ID AA59440 standard; DNA; 1516 BP.
XX
AC AA59440;
XX
DT 14-NOV-2000 (first entry)
XX
DE Nucleotide sequence of UDP-galactose 4-epimerase.
XX
KW UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
KW raffinose; stachyose; raffinose; stachyose; carbohydrate metabolism;
KW starch level; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 1..1059
FT /*tag= a
FT /product= "UDP-galactose 4-epimerase"
XX
XX WO200047755-A2.
XX
PD 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-US003453.
XX
XX 10-FEB-1999; 99US-0119588P.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
XX Rafalski JA, Thorpe CJ;
XX WPI: 2000-549152/50.
XX
XX P-PSDB; AAB07820.
XX
XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
XX cells to e.g. decrease cell wall constituents and increase starch
XX content.
XX
XX Claim 2; Page 52; 61pp; English.
XX
XX The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is
XX responsible for the interconversion of UDP-glucose and UDP-galactose. As
XX the activity of the enzyme appears to be particularly limiting to carbon
XX flux into the raffinose pathway, reduction of its activity should
XX decrease the levels of raffinose and stachyose in seeds. The
XX polynucleotide is useful to transform plant cells to enhance the degree

CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
CC the expression of UDP-galactose-4-epimerase within a cell enables
CC modification of plant carbohydrate metabolism. This allows growth of
CC grains with reduced cell wall constituents (fiber) and increased levels
CC of starch
XX
SQ Sequence 1516 BP; 374 A; 439 C; 472 G; 229 T; 0 U; 2 Other;
Query Match 36.3%; Score 476.4; DB 3; Length 1516;
Best Local Similarity 66.4%; Pred. No. 2.9e-112;
Matches 684; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
Qy 54 CATCTCTGTCACCGTGGTGGTTCATTGGCACCACACCGTCTGTCAGCTTCTCAA 113
Db 12 CGTCTGGTACCGCGCGCGGGTTTCGCGACGACACCGTCTGTCAGCTGCTGGA 71
Qy 114 AGCTGGCTTTCAGCGTTTCAATAATCGACAATTCGATAACTCCGTCTATGAAGCAGTGA 173
Db 72 GAAGGGCTACCGCTGACCGCGCTCGACAACCTCCACAACCTCCGTCGCCGAGGCTCGA 131
Qy 174 CCGCTCGGCAAGTGGTGGCTCTGCTTCTCAGAACCTCAATTCACCCAGGCGA 233
Db 132 CCGCTCGGCAACATCGTCGCGCGCGCTCTCGCGCGCTCAATTCATCTTCGGGGA 191
Qy 234 TCTCCGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATGCCGTGAT 293
Db 192 CTTGACATGAGTATGCTGAGAGGCTCTTCGCGCCCAAGAGTACGACGCGGTGAT 251
Qy 294 CCATTTCTGGCTTGAAAGCGTGTGAAAGGTTGCGAAGCCCGCTGCTATTTTGA 353
Db 252 ACATCTCGCGGCTCAAGGCGGTGGCGGAGAGCGTGGCGCACCCCGAGATGTACAACCG 311
Qy 354 TTTTAATTTGGTGGCACCATCAACCTCTAGAGTTTATGCAAGTATATTTGCAAAA 413
Db 312 CAACAACATCTCGCACCGTCAACCTCTACAGCTCATGAGNAGCAGCGGTGCACAA 371
Qy 414 GATGGTTTCTCATCTGCAACCGTTTATGGCCAACTGCAAAAGATACCGTGTGAGGA 473
Db 372 GTTGGTGTCTCGTCTGCGGACCGTGTACGCGCAGCGGAGAGTGGCTCTCTCGA 431
Qy 474 GGATTTCAAGTTTACAAGCTATGATCCCTATGAGCGGACCAAGCTTTCTTGGAGAAAT 533
Db 432 GGATCTCCCCCTCAAGGCGCTCAACCGTACGCGAGGACCAAGCTGTACCTGGAGGAGAT 491
Qy 534 TGCCCGAGATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTAATGAGATATCTCAA 593
Db 492 GCTGCGGACTTACCAGCAGCGAACCCGAGTGGAGGACGATCTGCTGGCTACTTCAA 551
Qy 594 TCCAGTTGGGCTCATGAAGTGGCAACTCGGTGAGAGATCCCAAGGCGATCCCAATAA 653
Db 552 CCCATCGGCGCACACGAGAGCGCGACATCGGGGAGGACCCCAAGGCGCTCCCAACA 611
Qy 654 CCTCATGCTTTACATTCAGCAAGTAGCTGTTGGAAGATTGACTGAACCTCAATGTATACGG 713
Db 612 CTTGCTCCCTTACATCCAGCAGGTGGCGCTCGCCCGCGCGGCTCAACGTCTACGG 671
Qy 714 TCATGATTTCCAAAGAGGATGCTCTGCAATCGGAGCTATATCCATGTGATGAGCTT 773
Db 672 CCAGACTTACCGCACCCGCGACCGCGCTCAGGAGCTACATCCAGCTGGTGGACCT 731
Qy 774 GGCAGATGGCCATATTTGCTGCCCTCGCAAGAGCTCTTCAACAACGAGAACATAGGTGTGAC 833
Db 732 CGCCGACGCGCACATCGGGCGCTCGAGAGCTCTTCGCCACCCCTGACATCGCTGTGT 791
Qy 834 TGCTTACAACTCGGAACTGCTGCTGGAACATCTGTGCTTGAATGTTTACAGCTTGA 893
Db 792 GCGGTACAACTGGGGACCGGGCGCGGAGACGACCGTGTCTGGAGATGGTGGAGCGGTTCGA 851
Qy 894 AAAGCTTCTGGCAAGAAAATTTCCAGTAAATATTTGTCMAAGAGACCGGAGATGGCAC 953
Db 852 GAAGCATACGCGAAGAAAATTCGCGTGAAGATGTGCCCAAGAGGCGCGGCGATTCGA 911
Qy 954 TGAGTGTATGCTTACAGAGAGCTGAGAAAAGAACTTTGGTTGGAGGCAAACTATGG 1013

```

Db 912 GCGAGTGTACGGTCCACCGCCACAGGAGGAGTCTCGCTGGAGGGCCAGATGCGG 971
QY 1014 TGTGAGGAGATGTCCAGGACCAATGAATTTGGCAAGAAACAATCCCTGGGTTACGC 1073
Db 972 GATCAGGAGATGTCCAGGACCAAGTGGAACTGGGCCAAGAAGAACCGGTATGGCTACTG 1031
QY 1074 GGGGAAGCCT 1083
Db 1032 CGGCAACGCT 1041

RESULT 8
ADA69621
ID ADA69621 standard; DNA; 1065 BP.
XX
AC ADA69621;
XX
XX 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 2944.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX
XX Gen; ds.
XX
OS Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX
XX Claim 6; SEQ ID NO 2944; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
XX Sequence 1065 BP; 257 A; 298 C; 289 G; 221 T; 0 U; 0 Other;

Query Match 33.9%; Score 445; DB 7; Length 1065;
Best Local Similarity 66.1%; Pred. No. 3,1e-104;
Matches 675; Conservative 0; Mismatches 340; Indels 6; Gaps 2;

QY 55 ATTCTGGTCCAGCGTGGTCCGGTTTCATTTGGCACCCACAGCGTGTTCAGTCTTCAA 114
Db 25 ATCTCTGGTACGGCGCGCGGTACATCGCAGCCACACCGTCTCTCCAGTCTTCCAA 84
QY 115 GCTGGCTTCAGGTTTCAATATCGACAATTTCCGATACTCGTTCATGGAAGCAGTGGAC 174
Db 85 CTCGGCTTCGGGTTGTGGTCTCTCGACAACTCGACACGCTCCGAGCTCGCCATCTCTC 144

```

```

QY 175 CCGGTCCGCAAGTGGTTGGCCCTCTGCTTTCTCAGAACCTCCAATTACCCAGGGCGAT 234
Db 145 CCGGTCCAGGAACTCCCGG---ACACAAAGCCAAACACCTGACTTCGCAAGGTTGAC 201
QY 235 CTCGGAATAGGAGATGACTTTGGAGAACTCTTTCTCCAAAACAACTTTGATCCCGTGATC 294
Db 202 CTCGCGCACAAAGCGTTGGACCAATCTTCTCTCTCAAAGGTTTGAGGCTGTCTATC 261
QY 295 CACTTTGCTGGCTTGAAGCGGTTGCTGAAGCGTTGGAAGCGCGGTGCTATTTTGAT 354
Db 262 CATTTTGGCCGGCTGAAAGCTGTTGGCGAGAGCGTGCAGAGCCCTGCTTTACTACGAC 321
QY 355 TTTAATTTCGTTGGCAACCACTCAACCTCTACGAGTTTATGGCAAGTATATAATTCGAAAAG 414
Db 322 AACAACTCATCGGCACCATCACTCTCTCTGAGGTCATGCGCCACATGGTGCACCAAG 381
QY 415 ATGGTTTCTCATCTGCAACCGTTTATGGCCAACTGGAAGATACCGTGTGAGGAG 474
Db 382 CTGGTGTCTCATCTCCGAACTGTCTACGGGTGGCCCAAGGAGGTGCCCTGCACTGAA 441
QY 475 GATTTCAAGTTTCAAGCTATGAATCCTATGACGAGCAACCAAGCTTTTCTCTGAAAGAAAT 534
Db 442 GAATCCCACTTTGTGCAATGAACCCCTACGCGCAGACAAAGCTGGTATCGAGACATG 501
QY 535 GCGCGAGATATTCAGAAAGCTGAACAGAAATGGAAGATCATATTTACTGAGATCTTCAAT 594
Db 502 TGCGGGATCTGCATGCCTCAGACCCAACTGGAAGATCATACTGCTCCGATCTTCAAC 561
QY 595 CCAGTTGGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGCAATCCCAATAAC 654
Db 562 CTGTTTGGAGCTCACCCAAAGCGGTACATTTGTTGAGGACCCCTGCGGATCCCAACAC 621
QY 655 CTCANGCTTACATTCAGCAAGTACGCTGTTGAAAGATTGACTGAACTCAATGTATACGGT 714
Db 622 CTCATGCCCTTCGTCAGCAGGTGCTGTTGGCAGAGGCGCGCTTACCCTCTATGGA 681
QY 715 CATGATTATCCACAGAGGATGCTCTGCTCGATCCGGGACTATATCATGTGATGACTTG 774
Db 682 ACCGACTACAACACCAAGATGGAATCTGGGTTCTGACTATATCATGTTGTTGATCTA 741
QY 775 GCAGATGGCCATATTTCTGCTCCCTGCGAAAGCTCTTCAACA---ACGAGAACATAGTGT 831
Db 742 GCGGATGGTCAATATCCCGCGTTAAGAAAGCTCTATGAAGATCTCATAGAATAGGATGT 801
QY 832 ACTGCTTACAACCTGGAACTGCTGCGTGAACATCTGCTTGAATGTTTACAGCATTT 891
Db 802 GAGGTGTACAATCTGGGCACTGGAAGGGGACATCTGCTGGAATGTTGACGATTC 861
QY 892 GAAAAGGCTTCTGGCAAGAAAATTCAGTAAATATATGTCCAAGAAAGACCGGGAGATGCG 951
Db 862 GAGAAAGCTTCTGGAAGAAAATCCCGCTTGTATTGCTGGACGAAGGCTCGAGATGCC 921
QY 952 ACTGAGTTTATGCTATCTACAGAGAGAGCTGAGAAGAACTTGGTTGGAGGCAACTAT 1011
Db 922 GAGATCGTTTACGCTCAAACTGCCAAAGCTGAGAAGAACTGAAATGGAAGGCAAAATAC 981
QY 1012 GGTGTGAGGAGATGTGACGAGCAACCAATGGAATTTGGCAAGAAACAATCCCTGGGTTAC 1071
Db 982 GGGGTAGAGAGATGTGACAGGACCTGTGGAATTTGGCGAGCAGAACCCCTACGGGTAT 1041
QY 1072 G 1072
Db 1042 G 1042

RESULT 9
AAAS9442
ID AAAS9442 standard; DNA; 1498 BP.
XX
XX AAAS9442;
XX
XX 14-NOV-2000 (first entry)

```


XX Nucleotide sequence of UDP-galactose 4-epimerase.
 DE
 XX
 XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
 KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
 KW starch level; ss.
 XX
 XX
 XX Oryza sativa.
 XX
 XX
 XX Key Location/Qualifiers
 FH 106..1170
 FT CDS /tag= a
 FT /product= "UDP-galactose 4-epimerase"
 XX
 XX W0200047755-A2.
 PN
 XX
 XX 17-AUG-2000.
 PD
 XX
 XX 09-FEB-2000; 2000WO-US003453.
 PF
 XX
 XX 10-FEB-1999; 99JS-0119588P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Carlson TJ, Rader GM, Farnodu OO, Kinney AJ, Pearlstein RW;
 PI Rafalski JA, Thorpe CJ;
 XX WPI; 2000-549152/50.
 DR P-PSDB; AAB07822.
 XX
 XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
 PT calls to e.g. decrease cell wall constituents and increase starch
 PT content.
 XX
 XX Claim 2; Page 55-56; 61pp; English.
 PS
 XX The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is
 CC responsible for the interconversion of UDP-glucose and UDP-galactose. As
 CC the activity of the enzyme appears to be particularly limiting to carbon
 CC flux into the raffinoseaccharide pathway, reduction of its activity should
 CC decrease the levels of raffinose and stachyose in seeds. The
 CC polynucleotide is useful to transform plant cells to enhance the degree
 CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
 CC the expression of UDP-galactose-4-epimerase within a cell enables
 CC modification of plant carbohydrate metabolism. This allows growth of
 CC grains with reduced cell wall constituents (fiber) and increased levels
 CC of starch
 XX
 XX Sequence 1498 BP; 415 A; 389 C; 360 G; 334 T; 0 U; 0 Other;
 SQ
 Query Match 33.9%; Score 445; DB 3; Length 1498;
 Best Local Similarity 66.1%; Pred. No. 3.6e-104;
 Matches 675; Conservative 0; Mismatches 340; Indels 6; Gaps 2;
 QY 55 ATCTGTGTCACCGGTGTGGCGGTTTCATGCGACCCACACCGTCTCAGCTTCTCAA 114
 DB 130 ATCTGTGTCACCGGTGTGGCGGTTTCATGCGACCCACACCGTCTCAGCTTCTCAA 189
 QY 115 GCTGGTTCAGCGTTTCATGCGACCAATTCGATACTCCGTATGGAAGCAGTGGAC 174
 DB 190 CTGGCTTCGGGTGTGCTTCGACCACTCGACCAACCGCTCCGAGCTCGCCATCCTC 249
 QY 175 CGGTTCGCGCAAGTGTGGTGGCCCTCTGTTTCAGAACCTCCAAATCCACCGGCGAT 234
 DB 250 CGGTTCAGGAACTCGCGCGG---ACACACCGCAACAACTCGACTTCGCGAAGGTGAC 306
 QY 235 CTCGGATAGGATGACTTGGAGAACTCTCTCCAAACCAACATTTGATGCGCGTATC 294
 DB 307 CTCGGACCAAGCAAGGTTGGACCAATCTCTCTCAAGGTTGAGGCTGTCATC 366
 QY 295 CACTTTGCTGGCTTGAAGCGGTGTGTAAGCGTTGCGAAGCCCGTCTGCTATTTGAT 354
 DB 367 CATTTTCGGGCTGAAAGCTGTGTGGCGAGAGCGGTGCGAAGCCCGTCTGTTACTACGAC 426

QY 355 TTTAATTTGGTGGCACCATCAACCTCTACGAGTTATTGGCAAAGTATTAATTCGAAAAG 414
 DB 427 AACAACTCATCGCACCATCACTCTCTCTGAGGTTCATGCGCCGACATGCTGCACCAAG 486
 QY 415 ATGTTTTCATCATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAG 474
 DB 487 CTGGTGTCTCATCATCCGCACTGTCTACGGTGGCCCAAGGAGTGCCTTCACCTGAA 546
 QY 475 GATTTCAGTTTCAAGCTATGAATTCCTATGACGCGACCAAGCTTTTCCTGGAGAAAT 534
 DB 547 GATCCCACTTTGTGCAATGAACCCCTACGCGCAGAACAAAGCTGTGTAATTCGAGACATG 606
 QY 535 GCCCGAGATATTGAAAGCTGACCAAGATGAAGATCATATTTACTGAGATCTTCAAT 594
 DB 607 TCGCGGATCTGCATGCTCAGACCCAACTGGAAGATCATCTGCTCCGATCTTCAAC 666
 QY 595 CCAGTTGGGGCTCATGAAAGTGGCAAACTCGTGAAGATCCCAAGGCGATCCCAATAAAC 654
 DB 667 CTGTTGAGCTACCCCAAGCGGTACATTTGTTGAGGACCCCTCGGCGATCCCAACAAC 726
 QY 655 CTATGCTTACATTCAGCAAGTAGCTGTTGGAGATTGACTGAATCAATGTATACGGT 714
 DB 727 CTCATGCCCTTCGTCACAGCGTCTGTTGGCAGGAGCGCGCCCTTACCGCTATGGA 786
 QY 715 CATGATTATCCACGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTTG 774
 DB 787 ACCGACTACACACCAAGGATGGAACCTGGGTTGCTGACTATATCCATGTTGATCTA 846
 QY 775 GCAGATGGCATATTTGCTGCGCTCGGAAAGCTCTTACA---ACGGAGAACATAGGTTGT 831
 DB 847 GCGGATGTCATATCGCGCGTTAAGGAAGCTCTATGAAGATTCTGATAGAATAGGATGT 906
 QY 832 ACTGCTTACAACTGGGAACCTGGTCTGGAACATCTGCTGAAATGGTTTACAGCATTT 891
 DB 907 GAGGTGTACATCTGGCACTGGAAGGACATCTGTCTGGAATGTTTTCAGCATTC 966
 QY 892 GAAAGGCTTTCGCAAGAAATATCCAGTAAATATGTCGAAGAACCGGAGATGCG 951
 DB 967 GAGAAAGCTTCTGAAAGAAATATCCGCTTGTATTTCTGGACGAGGCGCTGGAGATGCC 1026
 QY 952 ACTGAGTTTATGCTATCAGAGAGAGCTGAGAAGAACTGTTGTTGGAAGCAACTAT 1011
 DB 1027 GAGATCGTTTACGCTCAAACTGCCAAAGCTGAGAGGACTGAATGGAAGGCAAAATAC 1086
 QY 1012 GGTGTGAGGAGATGTGCGAGGACCAATGGAATTTGGCAAGAACAAATCCCTGGGGTTAC 1071
 DB 1087 GGGTAGAGGAGATGTGCGAGGACCTCTGGAATTTGGCGAGCAAGAACCCCTACGGTAT 1146
 QY 1072 G 1072
 DB 1147 G 1147
 RESULT 10
 ID AAA59443 standard; DNA; 1532 BP.
 XX
 AC AAA59443;
 DT 14-NOV-2000 (first entry)
 XX
 DE Nucleotide sequence of UDP-galactose 4-epimerase.
 XX
 KW UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
 KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
 KW starch level; ss.
 XX
 OS Glycine max.
 XX
 XX Key Location/Qualifiers
 FH 193..1242
 FT CDS /tag= a
 FT

PR	23-APR-1999	99US-0130891P.
PR	28-APR-1999	99US-01314449P.
PR	30-APR-1999	99US-01320408P.
PR	30-APR-1999	99US-0132407P.
PR	04-MAY-1999	99US-0132484P.
PR	05-MAY-1999	99US-0132485P.
PR	06-MAY-1999	99US-0132486P.
PR	06-MAY-1999	99US-0132487P.
PR	11-MAY-1999	99US-0132863P.
PR	11-MAY-1999	99US-01324256P.
PR	14-MAY-1999	99US-0134218P.
PR	14-MAY-1999	99US-0134219P.
PR	14-MAY-1999	99US-0134221P.
PR	14-MAY-1999	99US-0134370P.
PR	18-MAY-1999	99US-01343768P.
PR	19-MAY-1999	99US-01343941P.
PR	20-MAY-1999	99US-0135124P.
PR	21-MAY-1999	99US-0135353P.
PR	24-MAY-1999	99US-0135629P.
PR	25-MAY-1999	99US-0136021P.
PR	27-MAY-1999	99US-0136392P.
PR	28-MAY-1999	99US-0136782P.
PR	01-JUN-1999	99US-0137222P.
PR	03-JUN-1999	99US-0137528P.
PR	04-JUN-1999	99US-0137502P.
PR	07-JUN-1999	99US-0137724P.
PR	08-JUN-1999	99US-0138094P.
PR	10-JUN-1999	99US-0138540P.
PR	10-JUN-1999	99US-0138847P.
PR	16-JUN-1999	99US-0139119P.
PR	16-JUN-1999	99US-0139452P.
PR	16-JUN-1999	99US-0139453P.
PR	17-JUN-1999	99US-0139492P.
PR	18-JUN-1999	99US-0139454P.
PR	18-JUN-1999	99US-0139455P.
PR	18-JUN-1999	99US-0139456P.
PR	18-JUN-1999	99US-0139457P.
PR	18-JUN-1999	99US-0139458P.
PR	18-JUN-1999	99US-0139459P.
PR	18-JUN-1999	99US-0139460P.
PR	18-JUN-1999	99US-0139461P.
PR	18-JUN-1999	99US-0139462P.
PR	18-JUN-1999	99US-0139463P.
PR	18-JUN-1999	99US-0139750P.
PR	18-JUN-1999	99US-0139763P.
PR	21-JUN-1999	99US-0139817P.
PR	22-JUN-1999	99US-0139899P.
PR	23-JUN-1999	99US-0140354P.
PR	23-JUN-1999	99US-0140695P.
PR	24-JUN-1999	99US-0140823P.
PR	28-JUN-1999	99US-0140991P.
PR	30-JUN-1999	99US-0141287P.
PR	01-JUL-1999	99US-0141842P.
PR	01-JUL-1999	99US-0142154P.
PR	02-JUL-1999	99US-0142055P.
PR	08-JUL-1999	99US-0142390P.
PR	08-JUL-1999	99US-0142603P.
PR	09-JUL-1999	99US-0142920P.
PR	12-JUL-1999	99US-0142977P.
PR	13-JUL-1999	99US-0143542P.
PR	15-JUL-1999	99US-0144005P.
PR	16-JUL-1999	99US-0144085P.
PR	19-JUL-1999	99US-0144331P.
PR	19-JUL-1999	99US-0144332P.
PR	19-JUL-1999	99US-0144333P.
PR	19-JUL-1999	99US-0144334P.
PR	19-JUL-1999	99US-0144335P.
PR	20-JUL-1999	99US-0144332P.
PR	20-JUL-1999	99US-0144332P.

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139459P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140391P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149446P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149923P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 08-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157855P.
PR 18-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 18-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159638P.
PR 20-JUL-1999;	99US-0144884P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0144814P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160767P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160815P.
PR 22-JUL-1999;	99US-0145192P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160989P.
PR 23-JUL-1999;	99US-0145224P.	PR 25-OCT-1999;	99US-0161404P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145918P.	PR 26-OCT-1999;	99US-0161359P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161360P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146386P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 02-AUG-1999;	99US-0146390P.	PR 28-OCT-1999;	99US-0161993P.
PR 03-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		

Query Match 32.7%; Score 428.4; DB 3; Length 1286;
Best Local Similarity 65.1%; Pred. No. 6.3e-100;
Matches 665; Conservative 0; Mismatches 351; Indels 6; Gaps 2;

QY	51	ACACATTCGTCACCGGTGTCGGTTCATTGGACCCACACCCGTCGTTCACTTCT	110
DB	90	AAACGTTCTAGTAACGCGGTGTCGATACATCGGTAGTCACACGGGTCTCACTTCT	149
QY	111	CAAGCTGGCTCAGCGTTTCAATAATCGCAATTCGATACTCCGTCATGGAAGCAGT	170

Db 150 TCTCGGTGTTACTCGTGTAGTTGTTGATAATCTCGATAATCTCCGCGGTCTCT 209
 QY 171 GACCGCGTCCGCAAGTGGTTGGCCCTCTGTTCTCAGAACCTCCAAATTCACCCAGG 230
 Db 210 TCAACGTGTGAAGAAA---CTGCGCGCTGAACACGGTGAACGCCCTCTCTCTCCACAGT 266
 QY 231 CGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACAAATTTGATGCCGT 290
 Db 267 GGATCTTCGAGACAGTCTGCTCTTGAGAAGATATTTCTCAGAAACAAAGTTGATGAGT 326
 QY 291 GATCCACTTGTGCTTGAAGCGTGTCTGAAGCGTTCGAAGCGTTCGATTCGATTT 350
 Db 327 GATACATTTGTGGAACCTTGAAGCGTGTCTGAAGCGTTCGAAGCGTTCGATTT 386
 QY 351 TGATTTTAATTTGGTTGGCACCATCAACCTCTACGAGTTTATGGCAAAAGTATAATTCGAA 410
 Db 387 TAATAATATCTGTTGTAATTTACTCTCTTTTGAAGTTATGCTCAACATGGCTGCA 446
 QY 411 AAGATGGTTTCTCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGA 470
 Db 447 AATCTTGTGTTTTCATCATCAGCACTGTGTATGCTCTCCAAAGAAAGTTCTCTTATTA 506
 QY 471 GGAGGATTTCAAGTTTCAAGCTATGATCCCTATGAGCGGACCAAGCTTTTCTCGAAGA 530
 Db 507 AGAGGATTTCCAAATTTCTGTTTGAACCCATATGAGCAACAAAGCTATTCATTGAGGA 566
 QY 531 AATGCGGAGATTTCAAGAACTGACACGAAATGGAAGATCATATTAATGAGATCTT 590
 Db 567 AATCTCGCGATGTATACGGTTCTGACCCGGAATGGAAGATATATATTTAGTATTT 626
 QY 591 CAATCCAGTTGGGCTCATGAAGTGCACAACTCGGTGAAGATCCAAAGGATCCCAAA 650
 Db 627 CAATCTGTTGGTGCACATCTTAGTGTGACATTTGGTGAAGACCTCTGTTATTCAAA 686
 QY 651 TAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAGATTTGACTGAACCTCAATGTATA 710
 Db 687 CAATCTCATGCTTGTGTACAAAGTCTGTTGTTAGGAGACCTCATCTCACTGTCTT 746
 QY 711 CGGTATGATTATCCACGAGGATGCTCTGATCCGGACTATATCCATGTATGATGA 770
 Db 747 TGAAGATGATTAACAATACAAAGATGAACAGGGTTGAGATTTACATTCATGTTATGA 806
 QY 771 CTTGGCAGATGGCCATATTTGCTGCCCTGCGAAAGCTTTTCACACGGAAGAACATAGTTG 830
 Db 807 TTTAGCTGATGGACACATAGCGCTCTACGTAAGC---TAGAAGATTTGCAAGTCGGTTG 863
 QY 831 TACTGCTTACACCTGGGAACCTGGTCTGGAACATCTGTCTGTAATGTTTACAGCAT 890
 Db 864 TGAAGTGTACATCTCGGAACAGGAATGGAACATCAGTTCTTGAAATGTTGATGCTTT 923
 QY 891 TGAAGAGGCTTCTGGCAAGAAAATTCAGTAAATTTATGTCCAAAGAACCGGAGATGC 950
 Db 924 TGAAGAGCATCCGGAAGAAATCCCTCTGTTGATCGCGGACGTCGACCGGAGATGC 983
 QY 951 GACTGAGTTTATGCACTACAGAGAGCTGGAAGAACTTGGTTGGAAGGCAACTA 1010
 Db 984 TGAAGTTGTATACGCTTCGACGGAAGAGCAGAAAGTGAATTTGAATTTGAAGGCCAAATA 1043
 QY 1011 TGGTTGGAGGAGATGTGAGGAGCAATTTGGAATTTGGCAAGAACAAATCCCTGGGGTTA 1070
 Db 1044 TGGATTTGAGAGATGTGTAGATTTTGTGGAATCTGGGCTAGTAATATCTTATGGATA 1103
 QY 1071 CG 1072
 Db 1104 TG 1105

RESULT 14
 ID AAA59435
 XX AAA59435 standard; DNA; 479 BP.
 AC
 XX AAA59435;

DT 14-NOV-2000 (first entry)
 XX Nucleotide sequence of UDP-galactose 4-epimerase.
 XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
 KW raffinose pathway; raffinose; stachyose; carbohydrate metabolism;
 XX starch level; ss.
 OS Glycine max.
 XX W0200047755-A2.
 XX 17-AUG-2000.
 XX 09-FEB-2000; 2000WO-US003453.
 XX 10-FEB-1999; 99US-0119588P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
 PI Rafalski JA, Thorpe CJ;
 XX WPI; 2000-549152/50.
 DR P-PSDB; AAB07815.
 XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
 PT cells to e.g. decrease cell wall constituents and increase starch
 PT content.
 XX Claim 2; Page 44; 61pp; English.
 PS The present sequence encodes a UDP-galactose 4-epimerase. The enzyme is
 CC responsible for the interconversion of UDP-glucose and UDP-galactose. As
 CC the activity of the enzyme appears to be particularly limiting to carbon
 CC flux into the raffinose pathway, reduction of its activity should
 CC decrease the levels of raffinose and stachyose in seeds. The
 CC polynucleotide is useful to transform plant cells to enhance the degree
 CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
 CC the expression of UDP-galactose-4-epimerase within a cell enables
 CC modification of plant carbohydrate metabolism. This allows growth of
 CC grains with reduced cell wall constituents (fiber) and increased levels
 CC of starch
 XX
 SQ Sequence 479 BP; 120 A; 115 C; 104 G; 125 T; 0 U; 15 Other;
 Query Match 32.4%; Score 424.8; DB 3; Length 479;
 Best Local Similarity 94.1%; Pred. No. 3.6e-99;
 Matches 443; Conservative 0; Mismatches 27; Indels 1; Gaps 1;
 QY 9 CACTTCTCTCCCTCTTATTGCAAGCATGGTGTCTTCTCCCAACACATTTGTCACCG 68
 Db 3 CAATTCTCTCCCTCNCNAATGCAAGTGGGNGTCTCTCCCAACANATTTGTCACCG 62
 QY 59 TGGTCCCGGTTTCATTGGCACCCACACCGTCGTTTCAGCTTCTCAAAGCTGGCTTCAGCGT 128
 Db 63 NGNGGGCGGTTTCATTGGNACCCACACCGTCGTTTCANCTTCTCAAAGCTGGCTTCAGCGT 122
 QY 129 TTCAATATCGCAATTTGCAATTCGATTCGTCATGGAAGCATGGAACGCGTCGCCCAAGT 186
 Db 123 TTCAATATCGCAATTTGCAATTCGATTCGTCATGGAAGCATGGAACGCGTCGCCCAAGT 182
 QY 189 GGTGGCGCTCTGCTTCTCAGAACCTCCAAATTCACCCAGCGCATCTCCGGAATAGGA 248
 Db 183 GGTGGCGCTCTGCTTCTCAGAACCTCCAAATTCACCCAGCGCATCTCCGGAATAGGA 242
 QY 249 TGACTTGGAGAACTCTTCTCCAAACAAATTTGATGCGCGTATCCACATTTGCTGCTT 309
 Db 243 TGACTTGGAGAACTCTTCTCCAAACAAATTTGATGCGCGTATCCACATTTGCTGCTT 302
 QY 309 G-AAAGCGGTTGCTGAAGCGTTGCGAAGCGCGCTGCTTATTTGATTTAAATTTGCTG 367
 Db 303 GAAAGCGGTTGCTGAAGCGTTGCGAAGCGCGCTGCTTATTTGATTTAAATTTGCTG 362



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:35:58 ; Search time 113 Seconds
(without alignments)
6443.324 Million cell updates/sec

Title: US-09-913-064A-13
Perfect score: 1312
Sequence: 1 gcagagccacttctctcc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCITUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261.8	20.0	100848	4	US-09-596-002-39
2	261.6	19.9	13206	4	US-08-961-527-33
3	258.6	19.7	1131	4	US-09-540-236-260
4	247.2	18.8	3140	4	US-08-961-527-158
5	244	18.6	1488	4	US-09-023-655-880
6	241.4	18.4	1830121	4	US-09-557-884-1
7	241.4	18.4	1830121	4	US-09-643-980A-1
8	240.8	18.4	1249	3	US-09-113-536-1
9	240.8	18.4	1249	4	US-09-624-183-1
10	240.8	18.4	1249	5	PCR-US95-05785-1
11	229	17.5	1038	4	US-09-543-681A-776
12	199	15.2	1092	4	US-09-489-039A-5024
13	158.2	12.1	307	4	US-09-313-294A-7524
14	146	11.1	4597	4	US-09-424-311-1
15	144.8	11.0	381	4	US-09-134-000C-2119
16	118	9.0	1026	4	US-09-134-000C-2833
17	114.8	8.8	990	4	US-09-107-532A-2282
18	106.4	8.1	1410	4	US-09-424-311-2
19	97.6	7.4	351	4	US-09-328-352-3204
20	85	6.5	1026	4	US-09-489-039A-916
21	64.2	4.9	1035	4	US-08-961-527-311
22	62.4	4.8	907	4	US-08-545-528D-1
23	53.8	4.1	580073	4	US-08-481-190-15
24	50.8	3.9	222	3	PCR-US93-00869-15
25	50.8	3.9	222	5	US-08-232-463-14
26	48.4	3.7	7218	1	US-08-718-964-1
27	48.4	3.7	2649	2	US-09-596-002-39

RESULT 1
US-09-596-002-39
; Sequence 39, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 39
; LENGTH: 100848
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 39
; PUBLIC INFORMATION:
US-09-596-002-39

Query Match 20.0%; Score 261.8; DB 4; Length 100848;
Best Local Similarity 54.7%; Pred. No. 4.7e-64;
Matches 571; Conservative 0; Mismatches 457; Indels 15; Gaps 2;

Qy 31 AGCATGGTCTCTCCCAACACATCTGGTCCCGGTGGTCCCGGTTTCATTGGCACC 90
Db 51531 AGCGAGCATTTATTATCCAAAAAATCTGGTAAACAGCGGAGGGGTTTATCGGCTCA 51590
Qy 91 CACACCGCTGGTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAATTCGAT 150
Db 51591 CATACCTTGGTGGCTGATACAGCGGGATTGTGCGAGTGGTGTATGATAATTTATCC 51650
Qy 151 AACTCCGTCAAGAGCAGTGGACCGGTCCCGCAAGTGGTGGCCCTCTGCTTTCTCAG 210
Db 51651 AATTCAAGTTTGTGTGAGTGCAGCGAGTTGAGCAATTTGTCGAACATATG----- 51705
Qy 211 AACCTCAATTCACCCAGGCGGATCTCCGAAATAGGATGACTTGGAGAACTTCTTCC 270
Db 51706 -----AATTATCAAGTGAATTTGGATAAACGCATCTTGACGCTGTTTTAAG 51758
Qy 271 AAAACAACATTGATGCCGTGATCCACTTGTGGTGAAGCGGTGCTGTAAGCGTT 330
Db 51759 GCATCATCAGTTTTTGTGTAGTTCATTTTCAGGGCTAAAGAGCTGTGGAGAAAGCACC 51818

ALIGNMENTS

```
QY 331 GCGAAGCCCGTCGCTATTTGATTTTAAATTTGGTTGGCACCATCAACCTCTACGAGTTT 390
Db 51819 AAAATCCCTTAAATATATCAAAATATATGTTGTCGGCACACTCAGTCTGTAGAGCTG 51878
QY 391 ATGGCAAGATATAATGCAAAAGATGTTTCTCATCATCTGCACCGTTATGGCCAA 450
Db 51879 ATGGCAAAATATGGGGTAAAAAATTCGGTTTTTTCATCATCGCAACCGTTATGGCAGC 51938
QY 451 CTGAAAAGATACCGTGTGAGGAGGATTTCAAGTTTACAAGCTATGAATCCCTATGGACGG 510
Db 51939 TCAATTCGCTACCAATCACCAGAGATATCGCGGTTTTTGTACAGTCTTATGGTCAA 51998
QY 511 ACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTCAGAAAGCTTCAGAACGATGAGAG 570
Db 51999 AGTAAATGATGGTATGAGCATATATGGAGGATTTGGTGAATGCTGTGATGATACATGGAAT 52058
QY 571 ATCATATTACTGAGTACTTCAATCCAGTTGGGGCTCATGAAGTGGCAAACTCGGTGAA 630
Db 52059 GTTGCTGTTTGGGTTATTTTAAATCCAAATGGTGACATGAAGATGGCGGAATGGTGAG 52118
QY 631 GATCCCAAGGGATCCCAATTAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAGA 690
Db 52119 GATCCAAACGATATTCCTCAATTAACCTGATGCCATACATCAGTCAAGTGGCGGTAGGTAA 52178
QY 691 TTGACTGAATCAATGTATACGCTCATGATTATCCAAAGGAGGATGCTCTGCGATCCGG 750
Db 52179 CTTAAACAGCTGCTGTATTTGGCAATGATATGAGACACCTGACGCGTACGGGGTGCGA 52238
QY 751 GACTATATCCATGTGATGACTTTGGCAGATGGCCATATTTGTCGCTCGGAAA---GCTC 807
Db 52239 GATTATATTATCATGTGCTGATTTGGCAAAAGGTCATGTGGCAGCACTGCATTTATTTGGTT 52298
QY 808 TTCACAAAGGAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 867
Db 52299 GGGCAATCTGTGGCATTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 52358
QY 868 GTGCTTGAATGTTACAGCTTTGAAAGGCTTTGCGCAAGAAATTTCCAGTAAATTA 927
Db 52359 GTTTTACAGCTTATTAAGCGTTTGGTCAAAATCTGTTCAAGCGTACCGTATGTTATT 52418
QY 928 TGTCGAAGAGACCGGAGATGCGTACGAGTTTATGCTATACAGAGAGAGCTGAGAAA 987
Db 52419 ACCAGACAGCAGCAGGTGACATTTGAGCGCTTATGCGAGTATGCGTATGAGTAAAGCCT 52478
QY 988 GAATTTGTTGGAAGCAAACTATGCTGTGAGGAGATGTGAGGAGCAATGGAATTTGG 1047
Db 52479 ATACTTAACTGGACACAGAGCTTGATATTGAGCGTATGCGTAGATACATGCGGGTGG 52538
QY 1048 GCAAGAACAAATCCCTGGGGTTA 1070
Db 52539 CAGTCAAAATATCCACACGGCTA 52561
```

RESULT 2

```
US-08-961-527-33/c
; Sequence 33, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13206 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-33
```

```
Query Match 19.9%; Score 261.6; DB 4; Length 13206;
Best Local Similarity 55.3%; Pred. No. 1.9e-64;
Matches 560; Conservative 0; Mismatches 434; Indels 18; Gaps 2;

QY 50 AACACATCTGTCACCGGTGGCGGTTTCATTTGSCACCCACACCGCTGTTTCAGCTTC 109
Db 9213 AAAAGATTTTGGTAACTGTTGGTGGCGGTTTATTTGGAACCCACACTGTTATTTGAATTGA 9154
QY 110 TCAAGCTGTCCTCAGCGTTTCAATTAATCGACAATTCGATACTCCGTCATGGAAGCAG 169
Db 9153 TCCAAGCAGGTATCAAGTTGTTGGTGGATAATCTTCTCAATAGCAATCGTAAGAGTT 9094
QY 170 TGAACCGCTCCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCCTCCAAATTCACCCAGG 229
Db 9093 TAGAAGTTTGTGAAGAAATCAGAGGAGTTGAGATTCCT-----TTCTATGAGG 9046
QY 230 GCATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCCAAAACAAACATTTGATGCGCG 289
Db 9045 CAGATATTCGTGACACAGATACCTCAGAGATATTTTCAAGCAGGAAGAAACGACTGGTG 8986
QY 290 TGATCCACTTTGCTGCTGGAAGCGTTGCTGGAAGCGTTGCGAGAGCCCGTCGCTATT 349
Db 8985 TCATTCACCTTTGCTGTTGAGGCGTTTGGCGAATCAACAGTATCCCTCTTGGCTACT 8926
QY 350 TTGATTTTAATTTTGGTTGGCACCATCAACCTCTACGAGTTTATGCGAAAGTATAATTGCA 409
Db 8925 ATGACACATATCGCTGGAACCTGTGAGCTTTTGAAGCCATGGAAGAAACAACTGTA 8866
QY 410 AAAAGATGTTTCTCATCATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTG 469
Db 8865 AAAACATCATCTTCAGTTCTTCTGCGACAGTTTACGGGGATCCGACACAGTCCCATCT 8806
QY 470 AGGAGGATTTCAAGTTTACAGCTATGAATCCCTTATGAGCGGACCAAGCTTTTCTGGAAG 529
Db 8805 TGGAGATTTCCACATTTTCAAGTACCAACCCATACCGTCTGTAAGCTCATGAGAGG 8746
QY 530 AATTTGCCGAGATATTCAGAAAGCTGAACAGATGGAAGATCATATTTACTGAGTACT 589
Db 8745 AATTTTGAATGATATTTTCAAAAGCAGACTCAGATGGAATGTTGTCTTCTGCTTACT 8686
QY 590 TCAATCCAGTTGGGCTCATGAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCCCA 649
Db 8685 TTAACCAATCGAGGCCATGAGTGGTGAATTTGGAGAAATCCAAACGGTATTTCCAA 8626
QY 650 ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATTTGACTGAATCAATGAT 709
Db 8625 ACAATCTTGGCATATGTGACTCAAGTAGCCGTTGGAAAAATTTAGAGCAAGTCAAGTGT 8566
QY 710 ACGTCAATGATTTCCAAAGGAGGATGGCTCTCCGATCCGGGACTATATCCCATGTGATG 769
Db 8565 TTGAGACGATTACGATACGGAAGATGGAACAGGTGTTGTTGATATATCCACGTTGTCG 8506
```

770	ACTTGGCAGATGGCATAATTCGTCCTCCCTCGGAAGCTCTTCAACAACGAGAACATAGTT	829
Qy		
Db		
8505	ATTGGCTAAGGTCACGTTTCGCGTTTGAATAAAATCCAA-----AAGGTTTCAGAC	8455
Qy		
830	GTACTGCTTACAACTGGGAACCTGTCGTGGAACATCTGTGCTTGAATGGTTACAGCAT	889
Qy		
8451	TAAACGTTTATAACCTTGGAACTGGTAAAGGTTACTCAGTTCITGAAATATCCAAACA	8399
Db		
890	TTGAAAAGGCTTCTGGCAAGAAATTCAGTAAAAATTATGTCCAGAGACCGGAGATG	949
Qy		
8391	TGAAAAAGCGTGGCGCGTCCATTCCCTACCGCATCGTAGAACGTCGCCAGGTGATA	8333
Db		
950	CGACTGAGGTTTTATGCATCTACAGAGAGCGTGAAGAAGAACTTGGTTGGAGGCCAACT	1005
Qy		
8331	TCGCTGCTGCTTACTCAGACCCGCCAAAGCTAAGCAGAACTCGGTTGGAGACAGAC	8277
Db		
1010	ATGTGTGGAGGAGATGTGCGAGGACCAATGGAAATTGGGCAAGAACAATCC	1061
Qy		
8271	TCGACATACCCAAATGTGTGAAGCAGCATGCGCTTTGGCAAGCAAGCATCC	8220
Db		

```

RESULT 3
US-09-540-236-260
; Sequence 260, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: GARY L. BRETON ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 260
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-260

```

Query Match	19.7%; Score 258.6; DB 4; Length 1131;
Best Local Similarity	54.6%; Pred. No. 3.7e-64;
Matches 569; Conservative	0; Mismatches 459; Indels 15; Gaps 2
QY	31 AGCATGGTCTCTCTCCCAACACATTCGTGTACCGGTGGTGCCGGTTTCATTGGCAAC 90
Db	
QY	91 AGCGAGCATATTATTACCAAAAATAATCTTGTAACAGCGAGGGGTATATCGCTCA 150
Db	
QY	91 CACACGGTTCAGCTTCTCAAAGCTGGCTTCAGCGTTTCAATAATACGACAATTTGAT 150
Db	
QY	151 CATACTTTGATTGAGCTGATACAGCGGGATTTGTGCCAGTGGTGATGATAATTTATCC 210
Db	
QY	151 AACTCCGTGATCGAAGCAGTGCACGGGTCCGCCAAGTGGTGC3CCCTGTGCTTCTTCAG 210
Db	
QY	211 AATTCAAGTTTTGTGTCAGTGACGGAGTGTAGCGAATTTGTGGCAACAATATG----- 265
Db	
QY	211 AACCTCCAAATCCACCGGGGGATCTCCGGAATAGGATGACATTGGAGAAACTCTTCTCC 270
Db	
QY	266 -----AATTTATTCAAGGTGATGTTTTGGATAAAACGCACTTGACGCTGTTTTTAAG 318
Db	
QY	271 AAAACAACATTTGATGCGGTGATCCACTTTGCTGGCTTGAAGGGTGTCTGAAAGCGTT 330
Db	
QY	319 GCACATCAGTTTTTTTGCTGTAGTTCATTTTCGAGGGCTAAAGCTGTGGAGAAAGCACC 378
Db	
QY	331 GCGAAGCCCGCTCGCTATTTTGAATTTAATTGGTTGGCCACCATCAACCTCTACGAGTTT 390
Db	
QY	379 AAAAAATCCCTTAAAAATATTATCAAAATAATGTTCTCGGCACACTCAGTCTGTTAGAGCTG 438
Db	
QY	391 ATGGCAAGATATAATTGCAAAAAGATGGTTTTTCTCATCTCTGCAACCGTTTATGGCCAA 450
Db	
QY	439 ATGGCAAAATATGGGGTAAAAAATGGCGTTTTTTTTCATCATCGGCAACCGTCTATGGCAGC 498
Db	
QY	451 CTTGAAAAGATACCGTGTGAGGAGGATTTCAAGTTACAAGCTATGAATCCCTATGGACGG 510
Db	

RESULT 4
US-08-961-527-158/c
; Sequence 158, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961.527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders

;	REGISTRATION NUMBER:	36,373	
;	REFERENCE/DOCKET NUMBER:	PB340P1	
;	TELECOMMUNICATION INFORMATION:		
;	TELEPHONE:	(301) 309-8504	
;	TELEFAX:	(301) 309-8512	
;	INFORMATION FOR SEQ ID NO:	158:	
;	SEQUENCE CHARACTERISTICS:		
;	LENGTH:	3140 base pairs	
;	TYPE:	nucleic acid	
;	STRANDEDNESS:	double	
;	TOPOLOGY:	linear	
;	US-08-961-527-158		

Query Match	18.8%;	Score 247.2;	DB 4;	Length 3140;
Best Local Similarity	55.0%;	Pred. No. 1.2e-60;		
Matches 539;	Conservative 0;	Mismatches 423;	Indels 18;	Gaps 2

QY	55	ATTCTGGTCACCGGTGGTCCCGGTTTCATTGGCACCCACACCGTCGTCAGCTTCTCTCAAA	114
Db	965	ATATTGGTAACAGCGGAGCTGGTTATATTGGTAGCCATACCGCTAGTAGAATTTACTAAAT	906
QY	115	GCTGGCTTCAGCGTTTCAATAATCGACAATTTGATCACTCCGTCATGGAACGACGTGGAC	174
Db	905	TTAGGAAGGAAGTTATCAITGTGCACAACCTTTCGAACCTCTAGCATCTTGGTGTCTAGAC	846
QY	175	CGCGTCCGCAAGTGGTTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAGGCGCAT	234
Db	845	CGTATTGAAGCAATTACAGGAATACGTC-----CGTGTTTACGAATTAGAT	798
QY	235	CTCCGGATAGGGATGACTTGGAGAACTCTTCTCCAAAACAACATTTGATCGCGTGATC	294
Db	797	GTTTGTGATAAACAGCAATTGAGAAAGGTATTTGAACAAGATTCGATCTGCTCAAT	738
QY	295	CACCTTGTGGCTTGAAGACGGTTCGCTGAAGCGTTTCGGAAGCCCGCTCGCTATTTTTGGAT	354
Db	737	CATTTTCAGGTTATAAAGCTGTCGGAGATCCGTCGACAAAGCCTCTGATGTACTACAAA	678
QY	355	TTTAAITTTGGTGGCACCATCAACCTCTACAGTTTATGCGAAAGTATTAATTTGCAAAAG	414
Db	677	AATAATATTATGAGTACATTCGGCACTTGTGTGAAGTGTATGTCAGAAATTTAATGTTAAAG	618
QY	415	ATGGTTTTCTCATCATCTGCAACCGTTTATGGCCAACTCGAAAGATACCGTGTGAGAG	474
Db	617	ATTGCTTTTCTTCAAGTGGCACTGTATATGGAAATTAACAATCAGTCACTCTAATTGNG	558
QY	475	GATTTCAAGTTACAAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTCGTGAAGAAATT	534
Db	557	ACGATGCAACAAAGTGCACAAATCTCTATGGGTATACGAAAGTGTATGCTTTCAGCAAAAT	498
QY	535	GCCCGAGATATTACAAAAGCTGAAACAGATGGAAGATCATATTACTGAGATACTTCAAT	594
Db	497	TTAAAGATGTCATGTGGCAGATTCAGATGGAGTATTGCGTGTGCTCGTTATTTCAT	438
QY	595	CCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAATTAAC	654
Db	437	CCAATTGGTGCTCATGAGTCTGGTTGATTGGGGAAGATCCCTTCAGGAATTCCTAACCAAC	378
QY	655	CTCATGCCCTTACATTACGCAAGTAGCTGTGTGGAGAGTTGACTGAACTCAATGTATACGGT	714
Db	377	TTGATGCCCTTATTATTTGCAAAATGACGGTAGGTAGCTATCAGAGCTAAGTGTGTTCGA	318
QY	715	CATGATTTACCAACGAGGGATGGCTCTGCCATCCGGACTATATCACTGTGATGCACTTG	774
Db	317	AATGATTTATGATACCCCTTGATGGTACTGGTGTGCGTGAATATATCCATCTAGTAGATTA	258
QY	775	GCAGATGCCCATTTTCTGCCCTCGGAAGAGCTCTTCACACCGGAGAACATAGTTGTACT	834
Db	257	GCATATGAGCATATAAAGAGCTTTAAGAAAAGT-----ATCTGAAAAAACAAGATTTAT	204
QY	835	GCTTCAACCTGGGAACCTGGTCGTGGAACTCTGTGCTTGAAATGGTTACAGCATTTGAA	894
Db	203	ATCTATAACCTAGGCTCGGAGAGGACCTAGTGTATTAACAATCTGTAAATACATTTGAA	144

Qy	895	AAGCTTCTCGCAGAGAAATTCAGTAAATTTATGTCCTAAGAGACCGGAGATGCGACT	954
Db	143	AGTGTAAATAGATCCCTATTCTTTATAAATAGTTCCAGAGCGTTTCAGGAGCGTTGCA	84
Qy	955	GAGGTTTATGCATCTACAGAGAGAGCGTGGAGAAAGAACTTGTTGGAGAGGCAAACTATGGT	1014
Db	83	ACTGTATGCAATGCAGACAAAGCATATAAGAAATTAATTTGGAGGACACAAATATCG	24
Qy	1015	GTGAGGAGATGTGCGGGA	1034
Db	23	ATTGAAGACATGTGTAGAGA	4
RESULT 5			
US-09-023-655-880			
; Sequence 880, Application US/09023655			
; Patent No. 8607879			
; GENERAL INFORMATION:			
; APPLICANT: Cocks, Benjamin G.			
; APPLICANT: Susan G. Stuart			
; APPLICANT: Jeffrey J. Seilhamer			
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE			
; TITLE OF INVENTION: EXPRESSION			
; NUMBER OF SEQUENCES: 1508			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.			
; STREET: 3174 PORTER DRIVE			
; CITY: PALO ALTO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/023,655			
; FILING DATE: HEREMITH			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Zeller, Karen J.			
; REGISTRATION NUMBER: 37,071			
; REFERENCE/DOCKET NUMBER: PA-0001 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (650) 855-0555			
; TELEFAX: (650) 845-4166			
; INFORMATION FOR SEQ ID NO: 880:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1488 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: GENEBANK			
; CLONE: g1119216			
US-09-023-655-880			
Query Match 18.6%; Score 244; DB 4; Length 1488;			
Best Local Similarity 54.3%; Pred. No. 6.6e-60;			
Matches 565; Conservative 0; Mismatches 460; Indels 15; Gaps 3;			
Qy	56	TTCTGCTACCGGTGTGTGCGGGTTTCATTGGACCCACACACCGTCGTTCAAGTCTCAAG	115
Db	107	TGCTGTAAACAGTGGGGTGGCTACATTGGCAGCCACACCGTCTCTGGAGG	166
Qy	116	CTGGCTTCACGCGTTTCAATAATCGAACATTTGATAACTCCCGTCATGGAAGCAGTGGACC	175
Db	167	CTGGCTACTTGGCTGGGTGATCGATCGAATTCATCAATGCTTCCGTCGAGGGGGTCCC	226

176 CGGT-----CCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAGG 229
Db TGCCTGAGAGCTTCGCGCGCTCCAGAGCTGACAGCCCGCTCTGTGGATTTGAGGAGA 286
Qy GCGATCTCCGGAATAGGATGACTTGGAGAACTCTCTCCAAAACAAATTTGATGCCG 289
Db TGGACATTTTGGACCGGAGCCCTACAGCGTCTCTCAAAAAGTACAGCTTTATGCGG 346
Qy TGATCCACTTTGCTGGCTTGAAGCGTTCGTAAGCGTTGGAGCGTGGAGCCCGCTGCTATT 349
Db TCATCCACTTTGCGGGCTCAAGCGCTGGCGAGTCGGTGCAGAAAGCTCTCGATTATT 406
Qy TTGATTTTAAATTTGTTGGACCATCAACTCTACGAGTTTATGGCAAGTATAATTGCA 409
Db ACAGATTTACCTGACCGGACCATCAAGCTTCTGGAGATCAATGAGGCCACCGGGTGA 466
Qy AAAAGATGGTTTCTCATCTATCTCAACCGCTTTATGGCCAACTGAAAGATACCGTGTG 469
Db AGAACCTGGTTCAGCAGCTCAGCCACTGTGTACGGGAACCCCGCTACCTGCCCTTG 526
Qy AGGAGG--ATTTCAGTTTACAGCTATGATCCCTATGACGAGCCACAGCTTTCTCTGG 526
Db ATGAGGCCACCCACCGGTGGTTGTACCAACCTTACGGCAAGTCCAGTTCTTCATCG 586
Qy AAGAAATTTGCCGAGATATTTCAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGAT 586
Db AGAAATGATCGGGACCTGTGCCAGCAGACAAAGCTTGGAAAGTGTCTGCTCGCT 646
Qy ACTTCATCCAGTTGGGGCTCATGAAGTGGCAACTCGGTGAAGATCCCAAGGGATCC 646
Db ATTTCAACCCACAGGTGCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
Qy CAATAACCTCATGCTTACATTCAGCAAGTGTGTGGAAGTGTGCAACTCAATG 706
Db CCAACACCTCATGCTTATGCTCTCCAGTGGCGATCGGGCGAGGCGCTGATG 766
Qy TATACGTCATGATATCCAAAGGAGTGGCTTCGATCCGAGTCCGAGTATATCCATGTA 766
Db TCTTTGGCAATGACTATGACACAGAGATGGCACAGGTGTCCGGGATATACATCATGTG 826
Qy TGGACTTGGCAGATGGCCATATTGCTCCCTGCGAAAGCTCTTCAACAGGAGAACATAG 826
Db TGGATCTGGCCAGGGCCCATTTGACAGCTTAAGGAGCT-----GAAAGACAGTGTG 880
Qy GTTGTACTGTCTTCAACCTGGGAACTGGTCTGTAACATCTGTGTTGAATGGTTACAG 886
Db GCTCCGGATCTCAACCTGGCAGCGGCACAGGCTATTCAAGTGTGCGAGATGGTCCAGG 940
Qy CATTGAAAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCCAAGAGACCGGAG 946
Db CTATGGAGAGGCTCTGGGAAGAGATCCCGTACAAAGTGTGGCGCGGGAAGGTG 1000
Qy ATGCGACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAA 1006
Db ATGTGGCAGCTGTTACGCCAACCCAGCTGCGCCCAAGAGGAGCTGGGTGGACAGCAG 1060
Qy ACTATGTGTGGAGAGATGTGAGGACCAATGGAATTTGGCAAGAACATCCCTGGG 1066
Db CCTTAGGCTGGACAGATGTGTGAGGATCTGTGGCGCTGGCAGAGAGAAATCTTCAG 1120
Qy GTTACGCGGGAAGCCCTTGA 1086
Db GCTTTGGCACGAAGCCTGA 1140

RESULT 6

US-09-557-884-1/c

; Sequence 1, Application US/09557884

; Patent No. 6506581

; GENERAL INFORMATION:

; APPLICANT: Fleischmann et al.

; TITLE OF INVENTION: The Nucleotide sequence of

the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 18.4%; Score 241.4; DB 4; Length 1830121;

Best Local Similarity 54.6%; Pred. No. 1.5e-57;

Matches 557; Conservative 0; Mismatches 442; Indels 21; Gaps 3;

Qy	54	CATTCTGCTCACCGGTGGTCCGGTTTCATTGGCACCCACACCGCTGCTTCACTCAAA	113
Db	378456	CATTTTAGTGGGTGGTGGCTGGTTATTCGGTTCTCACACAGTTGTAGATTATTAAA	378397
Qy	114	AGCTGGCTTCAGCGTTTCAATATCGACAATTCGATACTCCGTCATCGAAGCAGTGGA	173
Db	378396	TGTTGGCAAGAGGTGGTGTATTAGATAATCTTTGCAATTCATCGCCAAAATCCCTGA	378337
Qy	174	CGCGTCCGCCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCCAATTCACCCAGGCGA	233
Db	378336	CGCGTAAACAAATTCAGGCA-----AAGACAAAGTTTATGAGGCGA	378289
Qy	234	TCTCGGAATAGGATGACTTGGAGAACTCTCTCCAAAACAAATTTGATGCGGTGAT	293
Db	378288	TATTTTAGATCGTCTTGTGTACAAAATTTTTCAGAAAATGAGATTAACTCGGTAT	378229
Qy	294	CACTTTCCTGGCTGAAAGCGTTGCTGAAAGCGTTGCGAAGCCCGTCCCTATTTTGA	353
Db	378228	TCACTTTCAGGGTTAAAGCCGTAGGGGAAAGTTTCAAAAGCCGACAGAAATATTACAT	378169
Qy	354	TTTTAATTTGTTGGCACCCTCAACCTCTACAGTTTATGCAAGTATATTTGCAAAA	413
Db	378168	GAAATGTCGCTGGCACCTTGTATTATTCAGAAATGAAAAGACGAGGTGTTGGAA	378109
Qy	414	GATGTTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGA	473
Db	378108	CTTTGTTATTAGCTCATCTGCAACGGTTTACGGGAGCCAAAATTTATYCCAAATTACAGA	378049
Qy	474	GGATTTCAAGTTTACAAGCTATG---AATCCCTATGGAGCGGACCAAGCTTTTCTCTGGAAGA	530
Db	378048	GGATTTGTAAGTCGGCGGTACCAACCCCTTATGTTACATCTAAATATATGTTGAGCA	377989

QY 531 AATTCCCGGAGATATTCAGAAAGCTGAACCAAGTGAAGATCATATTTACTGAGATCTT 590
Db 377988 GATTTTACGGATACAGAAAGCGGAACCAAAATTTAGCATGATCTTCCGGTATTT 377929
QY 591 CAATCCAGTTGGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAAA 650
Db 377928 TAATCCAGTTGGGGCGCATGAAGTGGCTTGATTGGTGAAGATCCAAATGGTATTCCAAA 377869
QY 651 TAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGTGAAGTGAAGTGAAGTGA 710
Db 377868 TAATTTATTACCTTATATTAGCAAGTTGCTATCGGTAAATTAGCACAACCTTTCTGTATT 377809
QY 711 CGGTCAATGATTATCCCAACAGGAGTGGCTCTCGATCCGGGACTATATCCATGTGATGGA 770
Db 377808 TGGTAGTGATTACGACACTCATGATGGAACAGGTGGCTGATTATTTCAATGATGGA 377749
QY 771 CTGGGAGATGGCCATATTGCTGCTCGGAAAGCTCTTCAACAGGAGAACATAGGTG 830
Db 377748 TTTAGCTGTGGGGCATTTAAAAGCGCTTCAACGTATGAAA-----ATGACGCTGGCTT 377695
QY 831 TACTGCTTCAACCTCGGAACCTGGTCTGGAACATCTGTGCTGAAATGGTTACAGCAT 890
Db 377694 ACAATTTATACCTTGTGACGGCCATGGTTATTCCTGATTAGATATGTTAAAGCCTT 377635
QY 891 TGAAGAGCTTTCGCAAGAAATTCAGTAAATTTATGTCGAAGAACCGGGAGATGC 950
Db 377634 TGAAGAGCAACAATATTACAAATTCATATAAACTTTGAGAACCGCTCTGGCGATAT 377575
QY 951 GACTGAGGTTTATGATCTACAGAGAGCTGAGAAAGCTTGGTTGGAAGCAAACTA 1010
Db 377574 TGCACCTGCTATTCGATCTAGTTAGGGGCAAAAGCTTGGTTGGTGGGAGAG 377515
QY 1011 TGGTGTGGAGAGATGTGAGGAGCAATGGAATGGGCAAAAGCAATCCCTGGGGTTA 1070
Db 377514 TGGTCTTGAAGAAATGATGCAAGACAGTGGAACTGGCAGAAAAATATCCAAAAAGGTA 377455

RESULT 7

US-09-643-990A-1/c
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Owen White
Mark D. Adams
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: FBI86P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1930121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 18.4%; Score 241.4; DB 4; Length 1830121;
Best Local Similarity 54.6%; Pred. No. 1.5e-57;
Matches 557; Conservative 0; Mismatches 442; Indels 21; Gaps 3;
QY 54 CATTTCTGGTCACCGGTGGTCCCGGTTTCATTTGGCACCCACACCGTCGTTCCAGTTCTCAA 113
Db 378456 CATTTTAGTCACGGTGGCGCTGTTATATCGGTTCTCACACAGTTGTAGATTTATTTAA 378397
QY 114 AGCTGGCTTCAGCGTTTCAATATACGACATTTTCGTAACCTCCGTCATGGAACGATGGA 173
Db 378396 TGTGGCAAGAGGTGGTGTATTAGATAATCTTTGCAATTCATCGCCAAAATCCCTTGA 378337
QY 174 CCGGTCGCGCAAGTGGTTGGCCCTCTGCTTCTCAGAACCTCCAATTCACCCAGGCGCA 233
Db 378336 GCGCGTAAACAAATTTACAGGCA-----AAGAGCAAAAGTTTATGAGGCGCA 378289
QY 234 TCCTCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACCAACATTTGATCCGCTGAT 293
Db 378288 TATTTTAGATCGTCTTGTTCACAAAATTTTTCAGAAAAATTTTTCAGAAAAATGAGATTA 378229
QY 294 CCATTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGCGAAGCGTTGCGAAGCCCGCTATT 353
Db 378228 TCATTTGAGGGTTAAAGCGCTAGGGGAAAGTTTCAAAAGCGGACAGATATTACAT 378169
QY 354 TTTTAATTTGGTGGCAACATCAACCTCTACGAGTTTATGGCAAGATATTAATTCAAAAA 413
Db 378168 GAACAATGCTGCGTGGCAACCTTGTATTAAATCAAGAAATGAAAAAGCAGGTGTTGGA 378109
QY 414 GATGTTTCTCATCTCAACCGTTTATGCGCAACCTGAAAAAGATACCGTGTGAGGA 473
Db 378108 CTTTGATTTAGCTCATCTGCAACGGTTTACGCGGACCAAAATTTATYCCAATTACAGA 378049
QY 474 GGATTTCAAGTTCAAGCTATG---AATCCCTATGGACGGACCAAGCTTTTCTCTGGAAGA 530
Db 378048 GGATTTGGAAGTGGCGGTACAACCAACCTTATGTTGATATATATATGTTGTTGAGCA 377989
QY 531 AATTGCCCGAGATATTCAGAAAGCTGAACCAAGTGAAGATCATATTTACTGAGATACTT 590
Db 377988 GATTTTACGGATACAGCAAAAGCGGAACCAAAATTTAGCATGCTATCTTGGGTATTT 377929
QY 591 CAATCCAGTTGGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAAA 650
Db 377928 TAATCCAGTTGGGGCGCATGAAGTGGCTTGATTGGTGAAGATCCCAAAATGGTATTTCCAAA 377869
QY 651 TAACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGTGAAGTGAAGTGAAGTGA 710
Db 377868 TAATTTATTACCTTATATTAGCAAGTTGCTATCGGTAAATTAGCACAACCTTTCTGTATT 377809
QY 711 CGGTCAATGATTATCCCAACAGGAGTGGCTCTGCGATCCGGGACTATATCCATGTGATGGA 770
Db 377808 TGGTAGTGATTACGACACTCATGATGGAACAGGTGGCTGATTATTTATTTATGATGGA 377749
QY 771 CTGGGAGATGGCCATATTGCTGCTCGGAAAGCTCTTCAACAGGAGAACATAGGTG 830
Db 377748 TTTAGCTGTGGGGCATTTAAAAGCGCTTCAACGTATGAAA-----ATGACGCTGGCTT 377695
QY 831 TACTGCTTACAACTCGGGAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 890

Db 377694 ACATATTTATACCTTGGTACGGCCATGGTTATTCGGTATTAGATATGGTAAAGCCTT 377635
Qy 891 TGAAGAGGCTTTCGCAAGAAAATTCAGTAAATATATGTCCAAAGAACCGGGAGATGC 950
Db 377634 TGAAGAGGCAACAATATTACAAATTCATATAAACTTGTAGAACCGCGCTCTGGCGATAT 377575
Qy 951 GACTGAGGTTTATGCATCTACAGAGAGAGCTGAGAAAGNACTTGGTTGGAGGCAACTA 1010
Db 377574 TGCACCTTCTCTATTCGATCTCTAGTTTACCGCAAAAGAGCTTGGTTGGGCGAAGC 377515
Qy 1011 TGGTGTGGAGGAGATGTGAGGAGCAATATGGAATTTGGCAAGAAACAATCCCTGGGGTTA 1070
Db 377514 TGGTCTTGAAGAAATGATCAAGACACGTTGGAATGGCAGAAATAATCCAAAGGGTA 377455

RESULT 8

US-09-113-536-1

; Sequence 1, Application US/09113536

; Patent No. 6153739

; GENERAL INFORMATION:

; APPLICANT: JI, H., ET AL.

; TITLE OF INVENTION: HUMAN UDP GALACTOSE-4-EPIMERASE

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; FILING APPLICATION NUMBER: US/09/113,536

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/462,966

; FILING DATE: June 5, 1995

; APPLICATION NUMBER: PCT/US95/05785

; FILING DATE: 11 MAY 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-430

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1249 BASE PAIRS

; TYPE: NUCLEIC ACID

; STRANDEDNESS: SINGLE

; TOPOLOGY: LINEAR

; MOLECULE TYPE: cdna

US-09-113-536-1

Query Match 18.4%; Score 240.8; DB 3; Length 1249;

Best Local Similarity 54.1%; Pred. No 5e-59;

Matches 563; Conservative 0; Mismatches 462; Indels 15; Gaps 3;

Qy 56 TTCTGTACCGGTGTGGCGGTTTCATTGGCACCCACACCGGTGCTTACGTTCTCAAG 115

Db 106 TGGTGTAAACAGGTGGCGGTGGCTACATTGGCACCCACACCGGTGCTTGGAGCTGCTGGAG 165

Qy 116 CTGGCTTACCGGTTCAATAATCGAATTTGGATAACTCCGTCATGGAAGCAGTGGACC 175

Db 166 CTGGCTACTTGGCTGTGGTCACTAATTAATTCCTAATATGCTTCCGTGGAGGGGCTCCC 225

RESULT 9

US-09-624-183-1

; Sequence 1, Application US/09624183

; Patent No. 6451577

; GENERAL INFORMATION:

; APPLICANT: JI et al.

; TITLE OF INVENTION: HUMAN URIDINE DIPHOSPHATE GALACTOSE-4-EPIMERASE

Qy 176 GCGT-----CCGCCAAGTGGTTGGCCCTCTCTTCTTCAGAAACCTCCCAATTCACCCAGG 229
Db 226 TGCCTGAGACCTTGGCGGGTCCAGAGCTGACAGCCGCTCTGTGGAGTTTGAGGAGA 285
Qy 230 GCGATCTCCCGAATPAGGATGACTTGGAGAAACTCTTCTCCAAACAAACAATTTGATGCCG 289
Db 286 TGGACATTTTGGACACGAGGAGCCCTACAGGCTCTCTTCAAAAAGTACAGCTTTATGGCGG 345
Qy 290 TGATCCACTTTGTCTGCTTCAAGCGGTTCTCAAGCGTTCGAAGCGTTGGAGAGCCCGCTGCTATT 349
Db 346 TCATCCACTTTGGCGGGCTCAAGCGCTGGCGAGTGGTGGAGAGCCCTCTCGAATTATT 405
Qy 350 TTGATTTAAATTTGGTTGGCACCATCAACCTCTACGAGTTTATGGCAAAAGTATAAATTGCA 409
Db 406 ACAGAGTTTAACTGACCGCGGACCATCCAGCTTCTGGAGATCATGAAGGCCCAACCGGGTGA 465
Qy 410 AAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTTGAAGAGATACCGTGTG 469
Db 466 AGAACCTGGTTTTCAGCAGCTCAGCCACTGTACGGGAAACCCCCAGTAGTCTGCCCTTG 525
Qy 470 AGGAGG--ATTTCAGTTTACAAGCTATGAATCCCTATGACGAGACCAAGCTTTTCTCTGG 526
Db 526 ATGAGGCCACCCACCGGTTGTGTACCAACCTTACGCAAGTCCCAAGTCTTCTCATCG 585
Qy 527 AAGAAATTGCCCGAGATATTAGAAAGCTGAACCAAGATGGAAGATCATATTACTGAGAT 586
Db 586 AGGAAATGATCCGGGACCTGTGCGAGGACAGAACTTGGAAACGCGAGTCTGCTGCGCT 645
Qy 587 ACTTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCC 646
Db 646 ATTTCAACCCACAGGTGCCATGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
Qy 647 CAATAAACCTCATGCTTTTACATTGCAAGTAGTGTGTGGAAGATTGACTGAATCAATG 706
Db 706 CCAACAACTCATGCTTTATGTCTCCAGGTGGCGATCGGGCGACGCGGAGGCCCTGAATG 765
Qy 707 TATACGCTCATGATTATCCACGAGGAGTGGTCTGCGATCCGCGACTATATCCATGTGA 766
Db 766 TCCTTGGCAATGACTATGACAGAGATGGCACAGGTGTCGGGATTCATCCATGTG 825
Qy 767 TGGACTTGGCAGATGGCCATATTGCTGCCCTCGGAAAGCTCTTCAACACGAGAACATAG 826
Db 826 TGGATCTGGCAAGGGCCACATTGACGCTTAAAGAGCT-----GAAAGAACAGTGTG 879
Qy 827 GTTGTACTGTACAACTGGGAACTGGTCTGCGAATCATCTGCTTGAATGTTACAG 886
Db 880 GCTGCCGAGTCTACAACTGGGCAACGCGACAGGCTTATTCAGTCTCCAGTGTGCTCCAG 939
Qy 887 CATTTGAAAAGGCTTCTGGCAAGAAAATTCAGTAAATTTATGTCCAAGAACCCGGAG 946
Db 940 CTATGGAGAGGTCTCTGGGAGAGATCCCGTACAGGTGGTGGCACGCGGGAAGTGT 999
Qy 947 ATGCACTGAGTTTATGCTATCTACAGAGAGCTGGAAGAACTTGGTTGGNAGCAA 1006
Db 1000 ATGTGGCAGTCTGTACGCCAACCCAGTCTGGCCCCAAGAGGAGCTGGGTGGGACAG 1059
Qy 1007 ACTATGTTGGAGGAGATGTGAGGAGCAATGGAATTTGGGCAAAAGAAATCCCTGGG 1066
Db 1060 CTTAGGGCTGACAGAGTGTGTAGGATCTCTGGCGCTGGCAGAACAGATCTCTTCAG 1119
Qy 1067 GTTACGGGGGAAGCCTTGA 1086
Db 1120 GCTTTGGCAGCAAGCCTGA 1139

```

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
  ADDRESSEE: HUMAN GENOME SCIENCES, INC.
  STREET: 9410 KEY WEST AVE
  CITY: ROCKVILLE
  STATE: MD
  COUNTRY: US
  ZIP: 20850
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/624,183
  FILING DATE: 24-Jul-2000
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 09/113,536
  FILING DATE: JUL-10-98
  APPLICATION NUMBER: 08/462,966
  FILING DATE: JUN-05-95
  APPLICATION NUMBER: PCT/US95/05785
  FILING DATE: MAY-1195
ATTORNEY/AGENT INFORMATION:
  NAME: JONATHAN L. KLEIN
  REGISTRATION NUMBER: 41,119
  REFERENCE/DOCKET NUMBER: PF162C1D1
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 301-610-5790
  TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1249 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 93..1137
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-624-183-1

Query Match      18.4%; Score 240.8; DB 4; Length 1249;
Best Local Similarity 54.1%; Pred. No. 5e-59;
Matches 563; Conservative 0; Mismatches 462; Indels 15; Gaps 3;

QY 56 TTCTGGTCAACGGTGGTGGCGGTTTCATTGGCACACACACCGTCTGTCAGCTTCTCAAG 115
DB 106 TGTGTGTAAAGTGGGCTGGCTACATTGGCAGCACACGCTGCTGGAGCTGTCTGGAGG 165
QY 116 CTGGCTTCAACGGTTCATTAATCGACAATTCGATACTCCGTCATGGAAGCAGTGGACC 175
DB 166 CTGGCTTCACTGGCTGGTGCATGATAAATTCATTAATGCTTCGCTGGAGGGGCTCCC 225
QY 176 GCGT-----CGCCAGTGGTGGCCCTCTCTCAGAACCTCAATTCACCCAGG 229
DB 226 TGCTTGAGAGCTTGGCGCGGCTCCAGAGAGCTCACAGCGCTCTGTGGAGTTTGAGGAGA 285
QY 230 GCGATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCCAAAACAACATTTGATGCGG 289
DB 286 TGGACATTTTGGACACAGGAGCGCTACAGCGTCTCTTCAAAAGTACAGCTTTATGGCGG 345
QY 290 TGAATCATTGCTGGCTGGAAGCGTGTGTAAGCGTTCGGAAGCGCCGCTCGCTATT 349
DB 346 TCATCCACTTTGGGGGCTCAAGGCGCTGGCGAGTCCGCTGAGAAGCGCTCTGGATTATT 405
QY 350 TTGATTTAATTTGGTTGGCACCATCACTCTACGAGTTTATGGCAAGATATAATTCGA 409
DB 406 ACAGAGTTAACTGACCGGAGCACCATCCAGCTTCTGGAGATCATGAAGCCCGAGGGA 465
QY 410 AAAAGATGGTTTCTCATCATCTGTGCAACCGTTTATGGCCCAACCTGAAAGATACCGTGTG 469

```

```

DB 466 AGAACCTGGTGTTCAGCAGCTCAGCAGCTGTGTACGGGAACCCCAAGTACCTGCCCCCTTG 525
QY 470 AGGAGG---ATTTCAGAGTTACAAAGCTATGAATCCCTATATGAGCGGACCAAGCTTTTCTCG 526
DB 526 ATGAGGCCCAACCCACGGGTGTTGTACCAACCTTACGGCAAGTCCAAGTTCTTCTCATCG 585
QY 527 AAGAAATTCGCGAGATATTCAAGAGCTGAACCAAGATGGAAGATCATATTACTAGAT 586
DB 586 AAGAAATGATCCGGGACCTGTGCCAGGCGAGACAAGACTTGGAAACGCGAGTGTGCTCGCT 645
QY 587 ACTTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAAGTCCCAAGGGGATCC 646
DB 646 ATTTCAACCCACAGAGTGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
QY 647 CAAATACCTCATGCTTACATTCAGCAAGTAGCTGTTGGAAGATTGACTGACTCAATG 706
DB 706 CCAACAACTCATGCTTATGCTCCAGGTGGCGATCGGGGACCGGGAGGCCCTGAATG 765
QY 707 TATACGCTCATGATTATCCAAACGAGGAGTGGCTCTCGGATCCGGGACTATATCCATGTGA 766
DB 766 TCTTTGCAATGACTATGACACAGAGGATGGCACAGGTGTCCGGGATTATCATCATGTG 825
QY 767 TGAATTTGGCAGATGGCCATATTGCTGCGCTGCGGAAAGCTTTCACAAAGGAGAACATAG 826
DB 826 TGGATCTGCGCAAGGGCCCAATTGCAGCTTAAAGGAAGCT-----GAAAGAACAGTGTG 879
QY 827 GTTGTACTGCTTACAACTGGGAACTGCTGCTGGAACATCTGCTGTTGAAATGTTTACAG 886
DB 880 GCTGCGGATCTACAACTGGGCAACCCAGCTGTGCGGACAGGCTATTTCAGTGTGCGATG 939
QY 887 CATTTGAAAGGCTTTGCGCAAGAAATTCAGTAAATTAATGTCCTCAAGAGACCGGGAG 946
DB 940 CTATGGAAGGTCTCTGGAAGAGAGATCCCGTACAAAGTGTGTGGCAGCGCGGAAAGGTG 999
QY 947 ATCGACTGAGTATTATGATCTACAGAGAGAGCTGAGAAAGAACTTGGTTGGAAGCAA 1006
DB 1000 ATGTCGAGTCTGTTAGCCCAACCCAGCTGTGCGGACAGGCTGTGCGGTTGACAGCAG 1059
QY 1007 ACTATGTTGGAGGAGATGTGAGGACCAATGGAATTTGGGCAAGAAACAATCCCTGGG 1066
DB 1060 CTTAGGGCTGGACAGGATGTGTGAGGATCTCTGGCGCTGGCAGAGAGAGATCCTTCAG 1119
QY 1067 GTTACGCGGGGAAGCCTTGA 1086
DB 1120 GCTTTGGCAGCAAGCCTGA 1139

RESULT 10
PCT-US95-05785-1
; Sequence 1, Application PC/TUS9505785
; GENERAL INFORMATION:
; APPLICANT: J.I.H., ET AL.
; TITLE OF INVENTION: HUMAN UDP GALACTOSE-4-EPIMERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05785
; FILING DATE: Concurrently
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-305
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1249 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
PCT-US95-05785-1

Query Match 18.4%; Score 240.8; DB 5; Length 1249;
Best Local Similarity 54.1%; Pred. No. 5e-59;
Matches 563; Conservative 0; Mismatches 462; Indels 15; Gaps 3

QY 56 TTCTGTGTACCGGTGGTCCGGTTTCATTGTGCACCCACACCGCTCGTTAGCTTCTCAAG 115
DB 106 TCGTGGTAACAGTGGGCTGGCTACATTGGCAGCCACACGGTCTGGAGCTCTGGAGG 165
QY 116 CTGGCTTCAGCGTTTCAATATCGACAAATTCGNATACTCCGTTCATGGAGCAGTGGACC 175
DB 166 CTGGCTACTTGCCTGTGCTATCGATAACTTCCATAATGCTTCCGTGAGGGGGCTCC 225
QY 176 GCGT-----CGGCCAAGTGGTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAGG 229
DB 226 TGCCTGAGAGCCTCGCGGGTCCAGGAGCTGACAGGCCGCTCTGTGGAGTTGAGGAGA 285
QY 230 GCGATCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAAATTTGATGCGG 289
DB 286 TGACACATTTTGCACACAGGAGGCCCTACAGCGTCTCTTTCAAAAAGTACAGCTTTATGCGCG 345
QY 290 TGATCACATTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGCGAAGCGTTGCGAAGCCCGTGCCTATT 349
DB 346 TCATCCACTTTGCGGGGCTCAAGSCGTGGCGAGTCCGTGTCAGAAAGCCTCTGGATTATT 405
QY 350 TTGATTTAAATTGGTTGGCACCATCACTCTACGAGTTTATGGCAAGATATATTGCA 409
DB 406 ACAGAGTTAACTGACCGGGACCATCCAGCTTCTGGAGATCATGAAGGCCACCGGGGTGA 465
QY 410 AAAAGATGGTTTTCTCATCTGCAACCGTTTATGGCCAACTGCAAAAGATACCGTGTG 469
DB 466 AGAACCTGGTTTCAGCAGCTCAGCCACTGTGTACGGGAACCCCGAGTACTCTCCGCTTG 525
QY 470 AGAGG---ATTTCAGTTACAGGTATGAATCCCTATGGACGACCAAGCTTTTCTCGG 526
DB 526 ATGAGGCCACCCACGGGTGGTTGTGACCAACCTTACGGCAGTCCAGTTCTTCATCG 585
QY 527 AAGAAATTGCCCGAGATATTCAGAAAGCTGAACCAGAAATGGAAGATCATATTACTGAGAT 586
DB 586 AGGAATGATCCGGGACCTGTGCCAGGCAGACAAGACTTGGAAACGAGTCTCTCGCTG 645
QY 587 ACTTCAATCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAGATGCCCAAGGGCATCC 646
DB 646 ATTTCACCCCAACAGGTGCCCATGCTCTGGCTGCAATTGGTGAGGATCCCCAGGGCATAC 705
QY 647 CAAATAACCTCATGCTTACATTACAGCAAGTAGCTGTTTGGAAAGTTGCACTGAACTCAATG 706
DB 706 CCAACAACCTCATGCTTATGTCTCCAGGTGGCGATCGGGCGCAGGGAGCCCTGAAATG 765
QY 707 TATACGCTCATGTTATCAACAGAGGATGGCTCTGCGATCCGGGACTATATCATCATGTGA 766
DB 766 TCTTTGCAATGACTATGACACAGAGGATGGCACAGGTGTCCGGGTTTACATCATATGTCG 825
QY 767 TGGACTTGGCAGATGCCCATATTGCTGCCCTCGAAAGCTCTTCCAAACGGAGAACATAG 826
R26 TGGATCTGGCCAAAGGCCACATTGACGCTTTAAGAGACT-----GAAAGAACAGTGTG 879

QY 473 AGGATTTCAAGTTACAAG---CTATGATCCCTATGGACGACCAAGCTTTTCCTGGAAG 529
Db 434 AAGATCACTGTTGGTGGCACTACCAATCTTTATGGCACCGATATATATATGTCGAGC 493
QY 530 AAATGCCCGAGATATTGAGAAAGCTGAACAGAGATGAAGATCATATTAATCTGAGACT 589
Db 494 GTATTTTGTCTGATCTATATATGAGATGAGAGTTGCTCTATCAGTTTACTGGCTACT 553
QY 590 TCAATCCAGTTGGGCTCATGAAAGTGCAGAACTCGGTGAAGATCCCAAGGGCATCCAA 649
Db 554 TCAATCCAGTAGTGTCTATCTGCACTCTGGCACTATGGGGAGAGATCCCTAAGGGTATTCCTA 613
QY 650 ATAACTCATGCTTTACATTTACAGAAAGTAGCTGTTGGAAGATTGACTGAACCTCAATGTAT 709
Db 614 ATAAATTAACGCTTATATCTCGCAAGTTGCGATAGTCTGCTGTAAGAAAGTCGCGGTAT 673
QY 710 ACGGTATGATTTATCAACAGGAGGATGCTCTGCGATCGGGACTATATCCATCTGATGG 769
Db 674 TTGGTGATGATTAACCAACAAAGACGGTACTGTTGATCGTACTATATCCATGTTATGG 733
QY 770 ACTTGCGAGATGGCATATTGCTGCCCTCGGAAAGCTCTTCAACAGGAGAACATAGTT 829
Db 734 ATTTAGCTGATGTCATATTGCGGCTTAATGCGTT-----AGGAGAAAGCAGGGT 787
QY 830 GTACTGCTTACAACTGGGAATGCTGCTGGAACATCTGCTTGAATGCTTACAGCAT 889
Db 788 TACATATCTATAACCTTAGCACAGGAAATGGCACAGTGTGATTTGAATGATTTGAAGCAT 847
QY 890 TTGAAAGGCTTCTGCAAGAAATTTCCAGTAAATATGTCGAAGAAAGACCGGAGATG 949
Db 848 TTGCTAAACCAAGTGTGAAGCAATCCCTTATGAATTCAGCCAGTCTGCGCGGAGATA 907
QY 950 CGACTGAGTTTATGATCTTACAGAGAGCTGAGAAAGACTTGGTTGGAAGGCAACT 1009
Db 908 TCGCACATATGTTGTTCAAGTCTCTGCAAGAGCAAAAGACTTACACTGGAAGCCATTC 967
QY 1010 ATGGTGTGAGAGATGTCAGGACCAATGGAATTTGGCAAGAACCAATCCCTGGGGTT 1069
Db 968 GCTCTATTGATGATGAGCGCGATGCTTGGCGTTGGCAATTCACAAATCTTACGGTT 1027
QY 1070 A 1070
Db 1028 A 1028

RESULT 12

US-09-489-039A-6024
; Sequence 6024, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6024
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6024

Query Match 15.2%; Score 199; DB 4; Length 1092;
Best Local Similarity 52.0%; Pred. No. 4.4e-47;
Matches 532; Conservative 0; Mismatches 470; Indels 21; Gaps 3;

QY 56 TTCTGGTCAACCGGTGGTGGCGGTTTCAATGGCACCAACCGGTGTTCAAGTCTTCAAG 115
Db 83 TTCTGGTTACAGGTGGTGGCGGTTTCAATGGCAATCATACCTCGGTTCAACTGCTGAGC 142

QY 116 CTGGCTTCAGCGTTTCAATAATTCGAACAATTTTCGATAACTCCGTCATGGAAGCAGTGGACC 175
Db 143 AGGGACACAGAGTGGTGTGATCCCTCGCAAAATCTCTGCAACAGCAA-----GCGCA 190
QY 176 GGGTCGCGCAAGTGGTGGCGCTCTGCTTCTCAGAACCTCCAAATTCACCCAGGGCGATC 235
Db 191 GGGTGTGCGCGTGTGAAAGCTCTGGCGGCAAGAGCCACCTTTATTGAAGCGGATA 250
QY 236 TCCGAATAGGATGACTTTGGAGAAACTCTTCTCAAAAACAACATTTGATGCGGTGATCC 295
Db 251 TTGCTAACGAAGCGCTGATGACGGAGATCTCCAGATCACGCCATTTGAGGAGTGTATCC 310
QY 296 ACTTTCCTGGCTTGAAGCGGTTGCTGAAGCGTGTGGAAGCGGTCGCAAGCCCGTCTGCTATTGATT 355
Db 311 ACTTCGCGGGCTGAAAGCGGTCGGGGAGTCCGTCGCCAAGCGCTGGAATATTACGACA 370
QY 356 TTAATTTGGTTGGCACCATCAACCTCTACGAGTTTATGGCAAAAGTATATAATTCGAAAGA 415
Db 371 ATAACTCACCGGCACACTGAAATTAGTCTCTGCCATGCGCGCGCTGGCGTGAAGAACT 430
QY 416 TGGTTTCTCATCATCTGCAACGTTTATGGCCAACTGGAAGATACCGTGTGAGGAGG 475
Db 431 TTATCTTTAGCTCTCTCCGCTACCGTCTACGGGGATCAGCCGAAATTCCTGATGTCGAAA 490
QY 476 ATTTCAAGTTTACAAGCTATG---AATCCCTATGGACGACCAAGCTTTTCTCGGAAGAAA 532
Db 491 GCTTCCGACCGGCACCCCGCAAGCCCTACGGCAAAAGCAAGCTGTGATGGAGCAGA 550
QY 533 TTGCCCCAGATATTTCAGAAAGCTGAACAGAAATGGAAGATCATATTACTGAGATCTTCA 592
Db 551 TCCTGACCGACCTGCAAGAAAGCCAGCCGAGTGGAGCATTTGCCCTGCTGCGCTACTTCA 610
QY 593 ATCCAGTTGGGCTCATGAAAGTGGCAAACTCCGTTGAAGATCCCAAGGGCATCCCAATA 652
Db 611 ACCGTCGCGCGCCACCCCGTGGGGAGCATGGGGAGACCCCGAGGGATCCCGAACA 670
QY 653 ACCTCATGCTTACATTCAGCAAGTAGTGTGGAAGATGACTGAACTCAATGTATACG 712
Db 671 ACCTGATGCCCTACATCGCCAGGTGCGCGTTGGCGTTCGCGAATCGCTGGCCATTTTTCG 730
QY 713 GTCATGATTATCCAAAGAGGGATGGCTCTGCGATCGGGAATATATCCATGTGATGACT 772
Db 731 GTACGACTATCCGACCGAAGNTGGTACCGCGTGGCGATTTACATCCATGTCATGAGCC 790
QY 773 TGGCAGATGGCCATTTGCTGCCCTCGAAAAGCTTTTCAACGAGAAACATAGTTGTA 832
Db 791 TCGCCGATGTCAGCTCGCGCCATGGAATACT-----GGCTGCAAAAGCCCGGGTGC 844
QY 833 CTGCTTACAACTGGGAACTGGTCTGTTGGAACATCTGTTGAAATCGTTTACAGCATTTG 892
Db 845 ATATCTAACACCTTGGCGCGCGGTGCGGAGCAGCGTCTCGACGTAGTCAATGCTTCA 904
QY 893 AAAAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCGAAGAACCCGGGAGATCGGA 952
Db 905 GCAAGGCTCGGCAACACCCATTAACCTATTTCGGCGCGCGCGCGGATGGGACCTCC 964
QY 953 CTGAGGTTTATGATCTTACAGAGAGCTGAGAAAGACTTGGTTGGAAGGCAAACTATG 1012
Db 965 CGGCTTACTGGCGGATGCTGCCAAGGCCGACCGCGAGCTGAACCTGGCGGTAAAGCGCA 1024
QY 1013 GTGTGGAGGAGATGTGACGGAGCCAAATGGAATTTGGGCAAGAAACAATCCCTGGGGTTACG 1072
Db 1025 ACCTGGAGCAAAATGGCGCAGGACACCTGGCAGTCCCGTCACTCCGAGGGTTATC 1084
QY 1073 CGG 1075
Db 1085 CAG 1087

RESULT 13
US-09-313-294A-7524
; Sequence 7524, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalngudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 7524
LENGTH: 307
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700381880H1
US-09-313-294A-7524

Query Match 12.1%; Score 158.2; DB 4; Length 307;
Best Local Similarity 70.6%; Pred. No. 1.1e-35;
Matches 211; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 640 GGATCCCAATACCTCATGCTTACATTCAGCAAGTAGCTCTTGGAAAGATTGACTGAA 699
DB 8 GGAATCCCAACATCTTATGCCCTATGTTTCAGCAAGTTCCGTTTGTAGAGGCCAGCT 67
QY 700 CTCATATGATACGCTCATGATTATTCACACGAGGAGTGGCTCTCGCATCCGGGACTATATC 759
DB 68 CTRACAGTTTGTAGAAATGACTATGCAACAGAGATGGGACTGGGTCCGAGATTACATC 127
QY 760 CATGTGATGACTGGCAGATGAGCCATATTCCTGCCCTCGAAAGCTCTTCACAAAGGAG 819
DB 128 CATGTGTTGACCTTGTCTGACGACATATTTGCTGATGAGAGCTTTTGGAACTCT 187
QY 820 AACATAGTGTACTGTCTTACAACTCGGAAGTGGTCTGTGGAAACATCTGTCTTGAATG 879
DB 188 AGCATAGGTTGAAGCTGACAACTTGGAAAGAGAGTACATCTGTGCTGGAGATT 247
QY 880 GTTACAGATTGAAAGGCTCTTGGCAAGAAATTCAGTAAATATGTCACAAAG 938
DB 248 GTTAAAGCATTTGAGAAGGCTTCTGGGAAGAAATACCTCTGATTTTGTGTGAAGACG 306

RESULT 14
US-09-221-017B-1093/c
Sequence 1093, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA: PCT/AU98/01023
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mon-oy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1093:
SEQUENCE CHARACTERISTICS:
LENGTH: 4597 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc feature
LOCATION: 1...4597
US-09-221-017B-1093

Query Match 11.1%; Score 146; DB 4; Length 4597;
Best Local Similarity 51.2%; Pred. No. 1.4e-31;
Matches 397; Conservative 0; Mismatches 370; Indels 9; Gaps 2;

QY 281 TTGATGCCGTGATCATTCTGCTGCTTGAAGCGGTTGCTGAAAGCGTTGCGAAGCCCC 340
DB 4549 TAGCAGGGGCTCATTCTATTTGTCAGAGTAAGCCCGTAGGCGAATCGGTGAGAGCGCG 4490
QY 341 CTCGCTATTTTGAATTTTGAATTTGTTGGCACCATCTACAGTTTATGGCAAGT 400
DB 4489 TTTTGTACTACAGNACAAATCTCTCTCACTGTGTAATATCTCTGAACTGTGAGCGAT 4430
QY 401 ATAATTGCAAAAGATGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGA 460
DB 4429 TCGGCACCTCGTGGGATCGTATTTTCTCATCTCTGATCGGTACGACAGCGGAGGTAT 4370
QY 461 TACCGTGTGAGGAGATTTCAGTTAC---AAGCTATGAATCCCTATCGACGACCAAGC 517
DB 4369 TGCCCGTAAACGAGAGGCTCGGATACAGGAGCTCTCTCTCCCTATGCAATACGAAGC 4310
QY 518 TTTTCTCTGGAAGAAATTCGCGAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATAT 577
DB 4309 AATCAACGAGGAGATTATCCGCGATGCTATCCATGCGGAGCGGTACAAGGCTATAC 4250
QY 578 TACTGAGATATCTTCAATCCAGTTGGGCTCATGAAGTGGCAAACTCGGTGAAGATCCCA 637
DB 4249 TTCTGCGCTACTTCAATCCGATCGGGGCACACCCACGCGCATATCGGAGAGCTTCTTA 4190
QY 638 AGGGCATCCCAATAACCTCATGCTTACATTACAGCAAGTAGCTGTGGAAGATTGACTG 697
DB 4189 ACGGATGCCAGAGATTGATCCCTTATTTGACTCAACGCGACGACGATCAGAGCG 4130
QY 698 AACTCAATGATATACCGGTGATGATATCAAGAGGAGTGGCTCTCGATCCGGGACTATA 757
DB 4129 AACTGAGGCTCTTCGCGCATGACTACGACACGCGCGGCTTTGTATCCGGGACTATA 4070
QY 758 TCCATGTGATGACCTTGGCAGATGCCATATTTGCTGCCCTGCGAAAGCTTTTCACAAAG 817
DB 4069 TCTATGTGTGATTTGGCCAAAGCGCATGTGGCTGCCATCGAACGATGCTGATGAAG 4010
QY 818 AGAACATAGTTGTATC-----TGCTTCAACCTCGGAACTGGTCTGTGGAACATCTGTGC 871

34 ATGGTGTCTTCTCCCAACACATTTCTGGTCAACCGGTGGTGGCGGTTTCATTTGGCACCAC 93
97 AGGATGGCGTCAAGGGAACAATTTCTGGTAACTGGAGAGCTGGATTTCATCGGATCTCAC 156
94 ACCGTGTTTCAAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 153
157 ACCGTGTTTCAAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 216
154 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTTCTGCTTTCTCAGAAC 213
217 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTTCTGCTTTCTCAGAAC 276
214 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 273
277 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 336
274 ACAACATTTGATCCCGTGGATCCACTTTGCTGGCTTGAAGCGGT 317
337 ACCAAATTTGATCCCGTGGATCCACTTTGCTGGCTTGAAGCGGT 380

Search completed: March 22, 2004, 20:20:05
Job time : 127 secs

QY 34 ATGGTGTCTTCTCCCAACACATTTCTGGTCAACCGGTGGTGGCGGTTTCATTTGGCACCAC 93
Db 97 AGGATGGCGTCAAGGGAACAATTTCTGGTAACTGGAGAGCTGGATTTCATCGGATCTCAC 156
QY 94 ACCGTGTTTCAAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 153
Db 157 ACCGTGTTTCAAGCTTCTCAAGCTGCTTCAGCGTTTCAATTAATCGACAATTTCCGATAAC 216
QY 154 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTTCTGCTTTCTCAGAAC 213
Db 217 TCCGTATGGAAGCAGTGGACCGCGTCCGCAAGTGGTGGCGCTTCTGCTTTCTCAGAAC 276
QY 214 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 273
Db 277 CTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAA 336
QY 274 ACAACATTTGATCCCGTGGATCCACTTTGCTGGCTTGAAGCGGT 317
Db 337 ACCAAATTTGATCCCGTGGATCCACTTTGCTGGCTTGAAGCGGT 380

Search completed: March 22, 2004, 20:20:05
Job time : 127 secs

4009 AGAAGCGAGCGATTTCGTCGAAGTCTTTAATATAGGTACGGTAGGGCGGTGAGCGTGC 3950
872 TTGAATGTTTACAGCAATTTGAAGGCTTCTGCGCAAGAAATTCAGTAAATTTATGTC 931
3949 TGAGCTTATCCGTACCTTCGAACGGGTGACAGGCTTCTGCTTCTCATCGGATAGTAG 3890
932 CAAGAGACCGGAGATGCGACTGAGGTTTATGCTATCTACAGAGAGCTGAGAAAGAAC 991
3889 GAAGCGCGAAGCGGATATCGAACAGGTATGGCCGAGCGAAGAAAGCCAAATGAGGTGT 3830
992 TGGTTGGAAGCAACTATGTTGTTGAGGAGATGTGAGGACCAATGGAATTGG 1047
3829 TGGTTGGAAGCTTTGGAGTCTGTTGGAAGACACACTGCTGAGTGCATGGCGTTGG 3774

RESULT 15
US-09-424-311-1
; Sequence 1, Application US/09424311
; Patent No. 6372477
; GENERAL INFORMATION:
; APPLICANT: JORSBOE, Morten
; BRUNSTEDT, Janne
; PETERSEN, Steen Guldager
; TITLE OF INVENTION: CLONING OF UDP-GALACTOSE EPIMERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/424,311
; FILING DATE: 24-Jan-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/1B98/00886
; FILING DATE: 27-MAY-1998
; APPLICATION NUMBER: GB 9710991.2
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 078883/0114
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 85..381
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-424-311-1

Query Match 11.0%; Score 144.8; DB 4; Length 381;
Best Local Similarity 69.4%; Pred. No. 8.7e-32;
Matches 197; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 19:15:04 ; Search time 492 Seconds
(without alignments)
9864.409 Million cell updates/sec

Title: US-09-913-064A-13

Perfect score: 1312

Sequence: 1 gcacagccactctctccc.....aaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1279.4	97.5	1481	12	US-10-424-599-36584
2	1139	86.8	1424	12	US-10-425-114-8664
3	656.2	50.0	2934	12	US-10-424-599-67518
4	644.4	49.1	1056	9	US-09-938-842A-2357
5	644.4	49.1	1056	11	US-09-938-842A-2357
6	603.2	46.0	1488	14	US-10-060-275-1
7	601.8	45.9	1056	14	US-10-060-275-3
8	539.4	41.1	1474	12	US-10-425-114-9965
9	433	33.0	1575	12	US-10-425-114-29813
10	431.4	32.9	1566	12	US-10-424-599-89167
11	431.4	32.9	1622	12	US-10-425-114-29766
12	430.2	32.8	1497	12	US-10-425-114-5152
13	430.2	32.8	1531	12	US-10-425-114-1181
14	403.8	30.8	1200	12	US-10-425-114-22869
15	402.2	30.7	1729	12	US-10-425-114-3333

16	395.8	30.2	1661	12	US-10-425-114-21374
17	361.2	27.5	1266	12	US-10-425-114-15649
18	337.2	25.7	1850	12	US-10-424-599-89166
19	312	23.8	641	9	US-09-770-149-615
20	286.8	21.9	1334	12	US-10-425-114-3010
21	280.5	21.4	1017	12	US-10-425-114-8209
22	280.5	21.4	1213	12	US-10-424-599-89163
23	278.6	21.2	2647	12	US-10-425-114-31639
24	277.8	21.2	5793	10	US-09-869-855A-3
25	277.8	21.2	5943	10	US-09-869-855A-2
26	272.8	20.8	1061	12	US-10-424-599-6052
27	261.6	19.9	13206	12	US-10-158-844-33
28	248.5	18.9	545	15	US-10-393-840-606
29	247.2	18.8	3140	12	US-10-158-844-158
30	244	18.6	1047	9	US-09-318-271-9
31	244	18.6	1047	14	US-10-303-664A-17
32	244	18.6	1488	9	US-09-964-824A-97
33	244	18.6	1488	9	US-09-964-824A-547
34	244	18.6	1488	9	US-09-954-531-142
35	244	18.6	1488	9	US-09-954-531-360
36	244	18.6	1488	14	US-10-303-664A-16
37	241.4	18.4	1830121	14	US-10-329-960-1
38	241.4	18.4	1830121	15	US-10-329-960-1
39	240.8	18.4	1249	13	US-10-201-298-1
40	239.2	18.2	1053	13	US-10-108-605-64
41	231.6	17.7	874	12	US-10-425-114-7955
42	230	17.5	1730	12	US-10-424-599-89168
43	225.4	17.2	14483	14	US-10-216-289-3
44	215.2	16.4	578	15	US-10-393-840-600
45	211.8	16.1	1011	15	US-10-429-812-9

ALIGNMENTS

RESULT 1

US-10-424-599-36584
; Sequence 36584, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yungwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36584
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133037C.1
; US-10-424-599-36584

Query Match	97.5%	Score	1279.4	DB	12	Length	1481
Best Local Similarity	99.5%	Pred. No.	0				
Matches	1283	Conservative	0	Mismatches	6	Indels	0
Gaps	0						
QY	8	CCACTTCTCTCCCTCTCTATTGCGAGCATGGTCTCTCTCCCAACACATCTGTCACCG	67				
Db	111	CCACTTCTCTCCCTCTCTATTGCGAGCATGGTCTCTCTCCCAACACATCTGTCACCG	170				
QY	68	GTGGTCCCGTTTCATTTGGCACCACACCGTCGTTCTCAGCTTCTCAAAGTGGTTCAGCG	127				
Db	171	GTGGTCCCGTTTCATTTGGCACCACACCGTCGTTCTCAGCTTCTCAAAGTGGTTCAGCG	230				
QY	128	TTTCAATATCGACATTTTCGATTAATCGTATGAGCAGTGGACCGCGTCCGCAAG	187				
Db	231	TTTCAATATCGACATTTTCGATTAATCGTATGAGCAGTGGACCGCGTCCGCAAG	290				


```
QY 188 TGGTTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAGCGGATCTCCGGATAGG 247
Db 189 TGGTTGGCCCTCTGCTTTCTCAGAACCTCCAAATTCACCCAGCGGATCTCCGGATAGG 248
QY 248 ATGACTTGGAGAACTCTTTCTCCAAACCAACATTTGATCCCGTGCATCTTGTGGCT 307
Db 249 ATGACTTGGAGAACTCTTTCTCCAAACCAACATTTGATCCCGTGCATCTTGTGGCT 308
QY 308 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGTGCATTTTGAATTTAAATTTGGTTG 367
Db 309 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGTGCATTTTGAATTTAAATTTGGTTG 368
QY 368 GCACCACTCAACCTCTACGAGTTTATGGCAAGTATAATTCGAAAAGATGGTTTCTCAT 427
Db 369 GCACCACTCAACCTCTACGAGTTTATGGCAAGTATAATTCGAAAAGATGGTTTCTCAT 428
QY 428 CATCTGCAACCGTTTATGGCAACCTGAAAGATACCGTGTGAGAGATTTCAAGTTAC 487
Db 429 CATCTGCAACCGTTTATGGCAACCTGAAAGATACCGTGTGAGAGATTTCAAGTTAC 488
QY 488 AAGCTATGAATCCCTATGACCGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTC 547
Db 489 AAGCTATGAATCCCTATGACCGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTC 548
QY 548 AGAAGCTGAACAGAAATGGAAGATCATATTAATCTGAGATCTTCAATCAGTTGGGCTC 607
Db 549 AGAAGCTGAACAGAAATGGAAGATCATATTAATCTGAGATCTTCAATCAGTTGGGCTC 608
QY 608 ATGAAAGTGGCAACTCGGTGAAGATCCCAAGGGATCCCAATTAACCTCATGCCCTACA 667
Db 609 ATGAAAGTGGCAACTCGGTGAAGATCCCAAGGGATCCCAATTAACCTCATGCCCTACA 668
QY 668 TTCAGCAAGTAGCTGTGGAAGATTGACTGAATCAATGTATACCGTCAATGATTTCAA 727
Db 669 TTCAGCAAGTAGCTGTGGAAGATTGACTGAATCAATGTATACCGTCAATGATTTCAA 728
QY 728 CGAGGATGGCTCTGGATCCGGGACTATATCCATCTGATGACTTGGCAGATGGCCATA 787
Db 729 CGAGGATGGCTCTGGATCCGGGACTATATCCATCTGATGACTTGGCAGATGGCCATA 788
QY 788 TTGCTGCCCTGGAAAGCTCTTCAACCGGAGACATAGTTGTACTGTCTACAACTGG 847
Db 789 TTGCTGCCCTGGAAAGCTCTTCAACCGGAGACATAGTTGTACTGTCTACAACTGG 848
QY 848 GAACTGGTGTGGAACATCTGTGCTTGAATGGTTACAGCTTTGAAAGGCTTCTGCA 907
Db 849 GAACTGGTGTGGAACATCTGTGCTTGAATGGTTACAGCTTTGAAAGGCTTCTGCA 908
QY 908 AGAAATTCAGTAAATTTATGTCGAAGAACCGGGAGATCGGACTGAGGTTTATGCAT 967
Db 909 AGAAATTCAGTAAATTTATGTCGAAGAACCGGGAGATCGGACTGAGGTTTATGCAT 968
QY 968 CTACAGAGAGCTGAGAAGAACTTGTGTTGGAAGGCAAACTATGGTGTGAGAGATGT 1027
Db 969 CTACAGAGAGCTGAGAAGAACTTGTGTTGGAAGGCAAACTATGGTGTGAGAGATGT 1028
QY 1028 CGAGGACCAATGGAATTCGGCAAGAACCAATCCCTGGGTTTACCGGGAGCCCTGAA 1087
Db 1029 CGAGGACCAATGGAATTCGGCAAGAACCAATCCCTGGGTTTACCGGGAGCCCTGAA 1088
QY 1088 TTAGCTTGAGAAATATCTGCTCATCTACGAATGCTTTTCAATAAATAGGCACTCTTA 1147
Db 1089 TTAGCTTGAGAAATATCTGCTCATCTACGAATGCTTTTCAATAAATAGGCACTCTTA 1148
QY 1148 TATAGAAATCTTTTATGTTTATGATTTGTTTATGAGTTCGTTGTATAATCTTGACAA 1207
Db 1149 TATAGAAATCTTTTATGTTTATGATTTGTTTATGAGTTCGTTGTATAATCTTGACAA 1208
QY 1208 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGATTTTAAACAATACTTTAAATTA 1267
Db 1209 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGATTTTAAACAATACTTTAAATTA 1268
QY 1311 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGATTTTAAACAATACTTTAAATTA 1370
Db 1312 AAAATTTGGAGCAATTTCAAGAGTTAAAGCTATGATTTTAAACAATACTTTAAATTA 1371
```

```
QY 1268 GACTGGCCATTGATTTCATATTGAAAAA 1296
Db 1371 GACTGGCCATTGATTTCATATTGAAAAA 1399
```

RESULT 2

```
US-10-425-114-8664
; Sequence 8664, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8664
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700787631_FLI
US-10-425-114-8664
```

```
Query Match      86.8%; Score 1139; DB 12; Length 1424;
Best Local Similarity 94.0%; Pred. No. 2.6e-300;
Matches 1195; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
```

```
QY 26 ATTGACATGCTGTCTTCTCCCAACACATTTCTGTCACCGGTGTGGCGTTTCATTG 85
Db 73 ACTCCAAATGGTGACCTGTCGCAAAAGATTTCTGTTACCGCGGTGCGGTTTCATAG 132
QY 86 GCACCCACACCGTCTGTTTCTCAGCTTCTCAAAGCTGGTTCAGCGTTTCAATATCGCAAT 145
Db 133 GCACCCACACCGTCTGTTTCTCAGCTTCTCAAAGGAGGCTTCAAGTTTCAATATCGCAAT 192
QY 146 TCGATAACTCCGTTCATCGAAGCAGTGGACCGCGTCCGCCAAGTGGTTGGCCCTCTGTTT 205
Db 193 TCGATAACTCCGTTCATCGAAGCAGTGGACCGCGTCCGCCAAGTGGTTGGCCCTCTGTTT 252
QY 206 CTGAAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAACTCT 255
Db 253 CTGAAACCTCCGAAATTCACCCAGGCGATCTCAGGAATAGGATGACTTGGAGAACTCT 312
QY 266 TCTCCAAAACAAATTTGATGCGGTGATCCACTTTTCTGCTTGAAGCGGTTGCTGAA 325
Db 313 TCTCCAGAACAACTTTTGTATGCGGTGATCCACTTTTCTGCTTGAAGCGGTTGCTGAA 372
QY 326 GCGTTGGAAGCCCGTCTGCTATTTTGAATTTTAAATTTGTTGGTGGCACCACTCACTACG 385
Db 373 GCGTTTCCAGCCCGTCTGCTATTTTGAATTTTAAATTTTGTGCGGCACCATCACTCTACC 432
QY 386 AGTTTATGGCAAGTATAATTCGAAAGATGTTTCTCATCATCTGCAACCGTTTATG 445
Db 433 AGTTTATGGCAAGTATAATTCGAAAGATGTTTCTCATCATCTGCAACCGTTTATG 492
QY 446 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 505
Db 493 GCCAACTGAAAAGATACCGTGTGAGAGGATTTCAAGTTTCAAGTATGAATCCCTATG 552
QY 506 GACGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAGCTGAACCAAGAT 565
Db 553 GACGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAGCTGAACCAAGAT 612
QY 566 GGAAGATCATATTACTGAGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACTCG 625
Db 613 GGAAGATCATATTACTGAGGTAATCAATCCCGTTGGGCTCATGAAAGTGGTAAACTCG 672
```

626 GTCAAGATCCCAAGGCGATCCCAATAAACCCTCATGCTTACATTCAGCAAGTAGCTGTG 685
673 GTGAAGATCCCAAGGCGATCCCAATAAACCCTCATGCTTACATTCAGCAAGTAGCTGTG 732
686 GAAGATGATGTAACCTCAATGTATACCGGTATGATTCACACGAGGAGTGTCTGTGGA 745
733 GAAGATGCGGGAACCTCAATGTATATATGCGCATGATTCACCAAGGAGTGTCTGTGGA 792
746 TCCGGGACTATATCCCATGTATGAGGCTTGGCAGATGCGCATATGCTGCGCTGCCAAGC 805
793 TCCGGGACTATATCCCATGTATGAGGCTTGGCAGATGCGCATATGCTGCGCTGCCAAGC 852
806 TCTTCAACCGGAGACATAGTGTGTAAGTGTCTTACAACTGGGAACTGGTGTGGAACAT 865
853 TCTTCAACCGGAGACATAGTGTGTAAGTGTCTTACAACTGGGAACTGGTGTGGAACAT 912
866 CTGTGCTTGAATGGTTACAGCAATTTGAAAGGCTTCTGGCAAGAAATTCAGTAAAT 925
913 CGTGTCTTGAATGGTTACAGCAATTTGAAAGGCGTCTGGCAAGAAATTCAGTAAAT 972
926 TATGTCACAAAGACCGGAGATGCGACTGAGGTTTATGCATCTACAGAGAGCTGAGA 985
973 TATGTCACAAAGACCGGAGATGCGACTGAGGTTTATGCATCTACAGAGAGCTGAGA 1032
986 AGAACTTGGTGGAAAGGCAAACTATGCTGTGGAGGAGATGTGCGAGGACCAATGGAAT 1045
1033 AGAACTTGGTGGAAAGGCAAACTATGCTGTGGAGGAGATGTGCGAGGACCAATGGAAT 1092
1046 GGGCAAGAAACATCCCTGGGGTTACGCGGGAAGCTTGAATAGCTTGAAGAAATATAC 1105
1093 GGGCAAGAAACATCCCGAGGTTACCGGGAGGCTTACCGGGAGGCTTGAAGAAATATAC 1152
1106 TGCTCATCTAGAAATGCTTTCACAAATAGGAGCTCTCTTATAGAAATATCTTTATGT 1165
1153 TGCTCATCTAGAAATGCTTTCACAAATAGGAGCTCTCTTATAGAAATATCTTTATGT 1211
1166 TTGATGATTTGTTAGGAGCTTCTGTTGTAATCTTGAAGAAATTTGCGAGATTT 1225
1212 TGATGATTTGTTAGGAGCTTCTGTTGTAATCTTGAAGAAATTTGCGAGATTT 1271
1226 CAAGAAGTTAAAGCTATGTTTAAACAAATACTTTAAATAGCTGCGCAATGATTTGA 1285
1272 CAAGAAGTTAAAGCTATGTTTAAACAAATACTTTAAATAGCTGCGCAATGATTTGA 1331
1286 TATTGAAAGAA 1296
1332 TATTGAGAGAA 1342

RESULT 3
US-10-424-599-67518
; Sequence 67518, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 67518
; LENGTH: 2934
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_31983C.1
US-10-424-599-67518

Query Match 50.0%; Score 656.2; DB 12; Length 2934;
Best Local Similarity 76.2%; Pred. No. 3.7e-168;
Matches 821; Conservative 0; Mismatches 253; Indels 3; Gaps 1;
QY 29 GCAGCATGGTGTCTCTCCCAACACACATTCGTGTACCGGTGGTGGCGGTTTCATTCGCA 88
DB 1319 GGAGGATGGTGTCTCCCGATGTCAACGGTTCTGTGTGACGGAGGGCGAGGTTCATCGAT 1378
QY 89 CCCACACCGTGTCTCAGCTTCTCAAGCTGGCTTTCAGCGTTTCAATTAATCGACAAATTCG 148
DB 1379 CCCACACCGTGTCTCAGCTTCTGAAGCAGGCTTTCAGGTTTCCATCATCGACAACTTG 1438
QY 149 ATAATCTCGTCATGGAAGCAGTGGACCGGTGCGGCAAGTGGTGGCCCTCTGCTTCTC 208
DB 1439 ACAATTCCTCATCGAAAGCGGTTTCACAGGTTGCGCGCTGGTGGTCTCATCTTTCCA 1498
QY 209 AGAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAAATCTTCT 268
DB 1499 ACAACCTCACTTCTTCCAGGATCTCCGAAATGTTTCAGGATTTGGAGGCGAGTGTCT 1558
QY 269 CCAAAACAAATTTGATGCGGTGATCCATTTGCTGGCTTGAAGCGGTTGCTGAAAGCG 328
DB 1559 CGAAAGCAAGTTTGTATGCGGTGATCCACTTTGCTGGCTGAAAGGTTGCGCGAAAGCG 1618
QY 329 TTGCGAGCGCCGTCGCTATTTGATTTTAAATTTGGTTGGCACCATCAACTCTACAGT 388
DB 1619 TTGCAAGCGCCGCGGTTTATACGAACTCTAGTGGGACCTTAAACCTTTTGAAG 1678
QY 389 TTATGGCAAGTATAATTTGCAAAAGATGGTTTTCTCATCTCTGCAACCGTTTATGCC 448
DB 1679 CAATGGCTAAATACAAAGTATAAATAATGGTTATATCATCATCGCAACTGTTTATGGG 1738
QY 449 AACCTGAAAGATACCGTGT---GAGGAGATTTCAAGTTACAGCTATGAATCCCTATG 505
DB 1739 AAGCTGATAGTCCCATGTTGAGGAGGAGGTCATTTGCGAGCATGAATCCGATG 1798
QY 506 GACGGAACCAAGCTTTTCTCGGAAGAAATTCGCCAGATATTCAGAAAGCTGAAACAGAT 565
DB 1799 GAAGAACAAAGCTGTTCTGTAGAGAAATAGCCAGAGACATCCAGAGGCGGAGACAGAT 1858
QY 566 GGAAGATCATATTAATGAGATATCTCAATCCAGTTGGGCTCATGAAGTGGCAACTCG 625
DB 1859 GGAGGATCATTTCTGCTTCTGATCTTCAATCCGGTTGGTGGCCCGAGAGTGGACAGAT 1918
QY 626 GTGAAGATCCCAAGGCGATCCCAATAAACCCTCATGCTTACATTCAGCAAGTAGCTGTG 685
DB 1919 GGGAGATCCAGGGGAATCCCCATTAACCTCATGCTTACATTCACCAAGTGGCCGTTG 1978
QY 686 GAAGATGACTGAACCTCAATGTATACCGGTATGATTTCAATCCAAAGGAGTGGCTCTCGGA 745
DB 1979 GCAGATTCCTCAGCTCAATGTTTATGCTCATGACTATCCCACTAAAGATGGCAACCCGA 2038
QY 746 TCCGGGACTATATCCATGTATGAGCTTGGCAGATGGCCATATTTGCTGCGCTGGAAGC 805
DB 2039 TCCGGGACTATATCCATGTATGAGCTTGGCAGATGGCCATATTTGCTGCGCTGGAAGC 2098
QY 806 TCTTCAACCGGAGAACATAGTTGTACTGCTTACAACTGGGAACTGGTGTGGAACAT 865
DB 2099 TTTTTCACACAGACACCATCGTTGTAGTGTCTTACAACTTGGAACTGGGCGTGGCACT 2158
QY 866 CTGTGCTTGAATGGTTTACAGCATTTGAAAGGCTTCTGGCAAGAAATTTCCAGTAAAT 925
DB 2159 CCGTGTGTAATGGTTGCTGCTTTGAAAGAGCTTCCGGCAAGAAATTTCCATTAATAA 2218
QY 926 TATGTCACAAAGACCGGAGATGCGACTGAGGTTTATGATCTTACAGAGAGCTGAGA 985
DB 2219 TGTGTCACAGAGACCGGGGATGCTACTGCTGTATATGATCTACGGAACAAGCTGAGA 2278
QY 986 AAGAACTTGGTTGGAAGGCAAACTATGCTGTGGAGGAGATGTGCGAGGACCAATGGAAT 1045
DB 2279 AAGAACTTGGTTGGAAGGCAAACTATGCTGTAGAGGAAATGTGCGAGGATTTTGGAAAT 2338
QY 1046 GGGCAAGAAACAAATCCCTGGGGTTTACCGGGGGAAGCCTTGAATAGCTTGAAGAAATA 1102

Db 2339 GGACAGCAAAATCCATGGGATACCAAGGGAGCATTAAATGAAATGGGTTGTA 2395
|||||
RESULT 4
US-09-938-842A-2357
; Sequence 2357, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2357
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2357
Query Match 49.1%; Score 644.4; DB 9; Length 1056;
Best Local Similarity 76.3%; Pred. No. 3.1e-165;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
QY 49 CAACACATTCGGTCCCGGTGGTCCCGGTTTCATTGGCACCCACACCGTTCAGTTCAGCTT 108
DB 19 CAGACATTCGTGTACTGGTGGTCTGCTGCTTATCGGACGCATACGTGTTCACATTT 78
QY 109 CTCAAAGCTGGCTTCAGCGTTTCAATATCGACAAATTCGAAATTCGATCGATGAGCA 168
DB 79 CTCAAAGATGGTTTTAAGGTTTCGATCATCGATAATTTTGATAACTCTGTTATCGAAGCT 138
QY 169 GTGGACCGCTCCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAG 228
DB 139 GTTGATAGATTAGGAGCTTGGTCTGCTGATCTCTCAGAGAGCTCGACTTCATCTG 198
QY 229 GCGATCTCCGGAATAGGATGACTTTGAGAAATCTTTCTCCAAACAAATTTGATGCC 288
DB 199 GGTGATCTAAGAAACAAAGGGGACATTCGAGAAATCTATTCTCCAGCAGAGATTTGATGCT 258
QY 289 GTGATCCACTTTGCTGGCTTGAAGCGTTCGTAAGCGTTGCTGAAGCGTTCGAGCCCGCTGCTAT 348
DB 259 GTGATTCATTTGCGGGCTTAAAGCTGGTGGTGGAGTGTGAAACCCCTCGCGGCTAC 318
QY 349 TTGATTTTAAATTTGGTTGGCACCATCAACCTCTACGATTTATGGCAAGTAAATTCG 408
DB 319 TTGACAATAACTTGGTTGGAACAATCAATCTATATGAGACCATGGCAAGTACAACATGC 378
QY 409 AAAAGATGGTTTCTCATCTCCACCGTTTATGSCCAACCTGAAAGATACCGTGT 468
DB 379 AAAATGATGGTGTTCATCTCTCCACTGTTTATGGACAACCTGAAAGATTTCCATGC 438
QY 469 GAGGAGGATTTCAAGTTCAAGCTATGAATCCCTATGACCGGACCAAGCTTTTCTCGAA 528
DB 439 ATGGAAGACTTTGAATTAAGGCTATGATCCTTATGTCGTACTAGCTCTTCTTTGAA 498
QY 529 GAAATTCGCGGATATTCAGAAAGCTGAAACAGAAATGGAAGATCATATTTACTGAGATAC 588
DB 499 GAAATAGCTAGATATTCAGAAAGCTGAAACCGGAATGGAATTTTCTGCTGAGGTAC 558
QY 589 TTCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGGATCCCA 648
|||||

Db 559 TTCAATCCTGTAGGAGCACATGAGAGTGCGAGTATTGGTGAGGATCCAAAGGCAATCCCC 618
QY 649 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATTGACTGAACTCAATGTA 708
DB 619 AATAACCTCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATTGACTGAACTCAATGTC 678
QY 709 TACGGTCAATGATTATCCAAACGAGGATGGCTCTGGGATCCGGGACTATATCCATGTGATG 768
DB 679 TATGGACATGACTATCCCAACCGGAGTGGTAGTGGGTAAGAGACTACATCCATGTGATG 738
QY 769 GACTTGGAGATGGCCATATTTGCTGCCCTCGCAAGAGCTCTTCACAAAGGAGAACATAGT 828
DB 739 GATTTAGCAGATGGCCATATTCGCTCGCTCAGGAAGCTATTTGCTGATCCAAAGATTTGT 798
QY 829 TGTACTGCTTCAACCTCGGAACTGGTCTGTGAAACATCTGTCTTCAATGGTTACAGCA 888
DB 799 TGTACTGCTTCAACCTAGGAGCTGGTCAAGGAAAGCTGTGTGTAGGAATGTTGCAGCT 858
QY 889 TTTGAAAAGGCTTTCTGGCAAGAAAATTCAGTAAATATATGTCCAAAGAGACCGGAGAT 948
DB 859 TTTGAAAAGGCTTTCCGGCAAGAAAATTCAGTAAAGCTGTCTCCGAGAGGTCAGGAGAT 918
QY 949 GCGACTGAGGTTTATGCTATCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAAAC 1008
DB 919 GCAACAGCAGTTTATGCTTCAACAGAGAGGCTGAGAAAGAACTTGGCTGGAGGCAAAA 978
QY 1009 TATGTTGTGGAGGAGATGTGAGGAGACCAATGGAATTTGGGCAAGAAACATCCCTGGGT 1068
DB 979 TATGAGTGGATGAGATGTGAGAGATCAGTGGAAATGGGCAACAAATAATCCATGGGT 1038
QY 1069 TACGCGGGAAGCTTGA 1086
DB 1039 TACCAGATAAGCTTTGA 1056
RESULT 5
US-09-938-842A-2357
; Sequence 2357, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2357
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2357
Query Match 49.1%; Score 644.4; DB 11; Length 1056;
Best Local Similarity 76.3%; Pred. No. 3.1e-165;
Matches 792; Conservative 0; Mismatches 246; Indels 0; Gaps 0;
QY 49 CAACACATTCGGTCCCGGTGGTCCCGGTTTCATTGGCACCCACACCGTTCAGCTT 108
DB 19 CAGACATTCGTGTACTGGTGGTCTGCTGCTTATCGGACGCATACGTGTTCACATTT 78
QY 109 CTCAAAGCTGGCTTCAGCGTTTCAATATTCGACAAATTTGGAATACTCCGTATGGAAGCA 168
DB 79 CTCAAAGATGGTTTTAAGGTTTCGATCATCGATAATTTTGATAACTCTGTTATCGAAGCT 138

QY 830 GTACTGCTTACAACTGGGAAGTGGTGGTGGAAATCTGTGCTTGAATGGTTACAGCAT 889
Db |||||
QY 942 GTGTTGCCCTACAAATTTGGGTACTGAAAGGCAAAATCTGTCTAGAGATGGTCTGCTT 1001
Db |||||
QY 890 TTGAAAAGGCTTCTGGCAAGAAATTCAGATTAATATATGTCCTCAAGAGACCGGGAGATG 949
Db |||||
QY 1002 TTGAAAAGGCTTCTGGCAAGAAATTCAGATTAATATATGTCCTCAAGAGACCGGGAGATG 1061
Db |||||
QY 950 CCACTGAGTTTATGCTTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACT 1009
Db |||||
QY 1062 CCACTGCTGTTTATGCTTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACT 1121
Db |||||
QY 1010 ATGGTGTGAGAGAGATGTCAGGACCAATGAAATGGGCAAGAACTTGGTTGGAGGCAAACT 1069
Db |||||
QY 1122 ATGGTGTGAGAGAGATGTCAGGACCAATGAAATGGGCAAGAACTTGGTTGGAGGCAAACT 1181
Db |||||
QY 1070 ACGGGGGAAGCTTGAATAGCTTGAAGAAATATAC 1105
Db |||||
QY 1182 ACCAATCAAGGCTTGAATAGCTTGAAGAAATATAC 1217
Db |||||

RESULT 7
US-10-060-275-3/c
; Sequence 3, Application US/10060275
; Publication No. US20030073828A1
; GENERAL INFORMATION:
; APPLICANT: DAI, ZIYU
; APPLICANT: SHI, LIFANG
; APPLICANT: HOOKER, BRIAN S.
; TITLE OF INVENTION: EPIMERASE GENE AND USE THEREOF
; FILE REFERENCE: 059440-0143
; CURRENT APPLICATION NUMBER: US/10/060,275
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/265,311
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Solanum tuberosum
US-10-060-275-3

Query Match 45.9%; Score 601.8; DB 14; Length 1056;
Best Local Similarity 73.8%; Pred. No. 1.4e-153;
Matches 765; Conservative 0; Mismatches 272; Indels 0; Gaps 0;

QY 50 AACACATTCCTGGTCAACCGTGGTGGCGGTTTCATTGGCAACCCACACCGCTCGTTCAAGCTTC 109
Db |||||
QY 1037 AAAATATTTGGTACTGAGGAGCTGGTTTCATTGGACACACACTGTGTGCAAGTTAC 978
Db |||||
QY 110 TCAAGCTGGTTCAGCGTTTCAATATCGACAATTCGATTAATCTCCGTCATGGAGCAG 169
Db |||||
QY 977 TGAATGAAGGGTTCAAGTTACCATCAITGATAACTTTTCATAATCTGTGGAAGAAGCTG 918
Db |||||
QY 170 TGGACGCGTCCGCCAAGTGGTTGGCGCTTCTGCTTCTCAGAACTCCCAATTCACCCAGG 229
Db |||||
QY 917 TTGATAGATCAGAGAAATAGTTGGTCTCACTTTCAGAACTTGTGATTCATTTGG 858
Db |||||
QY 230 GCGATCCCGAATAGGATAGCTTGGAGAACTTCTCCAAACCAACATTTGATGCGG 289
Db |||||
QY 857 GTGATATTGAACAAAGATGACTTGGAGAGCTATTTCTAAGAAAGAGTTTGTGCTGTG 798
Db |||||
QY 290 TGTATCCACTTGGCTGGAAGCGTGGTGGAGCGTGGAGCGCCCGCTCGCTATT 349
Db |||||
QY 797 TGTGCTATTTGCTGGCACTTAAGCTTGGAGAGAGTGTGTTAGGCCCTTCTTACT 738
Db |||||
QY 350 TTGATTTTAAATTTGGTTGGCAACATCAACCTTACAGTTTATGGCAAGTAAATGCA 409
Db |||||
QY 737 TTGAGAACATCTGATTTGGATCAATTAATCTTGTATTACGTATGCGCAAGTAAATGTA 678
Db |||||
QY 410 AAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTG 469
Db |||||

Db 677 AGAAGTTGGTTTTTTCATCATCTGCGACAGTTTATGCTCAGCTGAAAAGTTCCCTGTG 618
QY 470 AGAGAGTTTCAAGTTTCAAGCTATGAATCCCTATGAGCGGACCAAGCTTTTCCCTGGAG 529
Db 617 TGGAGGATTTTGAATGAAGGCTATGAATCCCTATGAGTGGTGGAAACAAAGCTTATTTCTTGAG 558
QY 530 AAATTCGCCGAGATATTGAGAAAGCTGAAACAGAAATGGAAGATCATATTTACTTGAGATACT 589
Db 557 ATATTGCTCGGATATCCAGAGGAGATCAAGATGGAATATCATATCTGTTGAGGATTT 498
QY 590 TCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTGGTGGAGATCCCAAGGCGATCCCA 649
Db 497 TCAACCCAGTAGGAGCTCATGAAAGTGGCAAACTGGGGAAGATCCAAAGGCGATTTCCCA 438
QY 650 ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAGAGATTGACTGAATCATATGAT 709
Db 437 ACATCTTATGCTTACATTCAGCAAGTAGCTGTTGGTAGATTGCCAGAGTTGAAATGAT 378
QY 710 ACGGTCAATGATTTCCAGAGGAGTGGCTTGGAGATCCGGAGTATATATCCATGTGATGG 769
Db 377 ATGGCAACGACTACCTTACACCTGATGGTACCGGATACGAGATTATATATCCATGTTTGG 318
QY 770 ACTTGGCAGATGGCCATATTGCTGCCCTCGGAAAGCTCTTCAACCGGAGAACATAGGTT 829
Db 317 ATTAGCGGACGCTCATGTTGTTGCACTTCAGAGACTTCTAAGGCAGATCATTTAGGTT 258
QY 830 GTACTGCTTACAACTGGGAACTGGTGGCAACATCTGTGCTTGAATGGTTTACAGCAT 889
Db 257 GTGTTGCCCTACAAATTTGGGTACTTGGAAAGGCAAACTCTGCTTAGAGATGGTTGCTGCTT 198
QY 890 TTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATATGTCCTCAAGAGACCGGGAGATG 949
Db 197 TTGAAAAGGCTTCTGGCAAGAAATTCAGTAAATATGTCCTCAAGAGACCGGGAGATG 138
QY 950 CGACTGAGGTTTATGCTTCTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACT 1009
Db 137 CCACTGCTGTTTATGCTTCTACTGAAAAGCTGAGAAAGGAGCTCGGTTGGAAGGCAAAAT 78
QY 1010 ATGGTGTGAGAGATGTCAGGAGCAATGGAATGGGCAAGAACTTCCCTGGGTT 1069
Db 77 ATGGTATTAATGAGATGTCAGGAGCAAGTGGAAATGGGCAAGCAAAATCTTGGGTT 18
QY 1070 ACGGGGGAAGCTTGA 1086
Db 17 ACCAATCAAGGCTTGA 1

RESULT 8
US-10-425-114-9965
; Sequence 9965, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9965
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700889134_FLI
US-10-425-114-9965

Query Match 41.1%; Score 539.4; DB 12; Length 1474;
Best Local Similarity 67.1%; Pred. No. 1.9e-136; Indels 103; Gaps 3;
Matches 865; Conservative 0; Mismatches 321;

QY 29 GCAGCATGCTGCTCTCCCAACACATCTCTGTCACCGGTGTCGCGGTTTCATTCGCA 88
DB 89 GGAGGATGCTGCTCCCGATGTCAACGGTTCTGTCAGCGGAGCGGCGAGCTTCATCGAT 148
QY 89 CCACACCGCTGCTGAGCTTCTCAAGCTGCTTCAGCGTTTCAATAATCGACAACTTCG 148
DB 149 CCACACCGGTGCTGAGCTTCTGAAAGCGGCTTCAGGGTTTCCATCATCGAACCTTG 208
QY 149 ATAACTCCGTTCATGGAAGCAGTGGACCGGTCGCGCAAGTGGTGGCCCTCTGCTTCTC 208
DB 209 ACATTTCCCTCATCGAAGCGTTTCAGAGTTTCGCGCGCTGTCGCTCTCATCTTCCA 268
QY 209 AGAACCTCCAAATTCACCGAGCGGATCTCCGNAATAGGATGACITGGAGAACTTCTC 268
DB 269 ACAACCTCACCTTCTTCCACGAGATCTCCGCAATGTTGAGATTTGGAGGCGAGTGTCT 328
QY 269 CCAAAACAACA-----TTTGTGCGGTGATCCACTTTGCTGCTTGAAGCGG 316
DB 329 CGAAAGCAAACTGAAACAGGTTTGTGCGGTGATCCACTTTGCTGGCTGAAAGGTG 388
QY 317 TTGCTGAAAGCGTTGCGAAGCCCGTGGCTATTTTGAATTTTAAATTTGGTTGGCACCATCA 376
DB 389 TCGCGAAGCGTTGCAAGCCCGCGGTATTTACGACCAACATCTAGTGGGACTATAA 448
QY 377 ACCTTACGAGTTTATGCGCAAGTATAATTTGCAAAAGATGGTTTCTCATCATCTGCAA 436
DB 449 ACCTTTTGAAGCAATGGCTTAATACAGTGTAAAAAATTTGTTATATCATCATCGGCAA 508
QY 437 CCGTTTATGCGCAACCTGAAAGATACCGTGT---GAGGAGGATTTCAAGTTACAGCTA 493
DB 509 CTGTTTATGGGAAGCTGTAGAGTCCATGTGCGAGGAGGAGTGCATTTGCGAGGCA 568
QY 494 TGAATCCCTATGACGACCAAGCTTTCTGTAAGAAATTCGCGAGATATTCAGAAAG 553
DB 569 TGAATCCGTATGGAAGCAAAAGCTGTTCTGTAAGAAATAGCCAGACATCCAGAGG 628
QY 554 CTGAACCAAGATGGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGGGCTCATGAA 613
DB 629 CGAGACAGATGAGGATCATTTCTGCTGATACCTTCAATCCGTTGTCGCCACGAGA 688
QY 614 GTGCAAACTCGGTGAGATCCGAGGCAATCCAAATACCTTCAGCTTACATTCAGC 673
DB 689 GTGACAGATTTGGGGAAGATCCAAGGGAAATCCCAATACCTCATGCTTACATTCACC 748
QY 674 AAGTAGCTGTTGGAAGATTGCACTGAATCAATGTATACGCTCATGATTATCCAAACGAGG 733
DB 749 AAGTCGCGTTGGCAGATTTGCTCAGCTCAATGTTTATGTCATGACTATCCCACTAAG 808
QY 734 ATGGCTCTGC-----743
DB 809 ATGGCACCCCGGTCACTCACTCACTCTTGTCTTTTCAATTCCTTTTCTTCTT 868
QY 744 -----GATCGGAGCTATATCCATGTG 765
DB 869 CTTTCTCTTCTTCTAGCTTACAAATATATCTGAGATTCGAGACTATATCCATGTA 928
QY 766 ATGACTTTGGCAGATGCCATATTTGCTGCCCTCGAAAGCTTTCACAAACGAGAAACATA 825
DB 929 ATGACTTTGGCAGATGCCATATTTGCTGCCCTCGAAAGCTTTTTCACACAGACCATC 988
QY 826 GGTGTACTGCTTACAACTGGGAACTGTCGTCGGAACATCTGTGCTTGAATGTTTACA 885
DB 989 GGTGTAGTGCCTTACATCTTTGGAACTGGGCTGGGCAATCCGCTTGAATGTTTGTCT 1048
QY 886 GCATTTGAAAAGGCTTCTGGCAAGAAATTTCCAGTAAATATGTCCAAAGAACCCGGGA 945
DB 1049 GCTTTGAAAAGCTTCGGCAAGAAATTTCCATTAATAATGTCTCCAGAAACCCGGG 1108
QY 946 CATGGACTGAGGTTTATGATCTACAGAGAGCTGAGAAAGAACTTGTGTTGGAAGCA 1005

DB 1109 GATGCTACTGCTGTATATGATCTACGGAACAGGCTGAGAAAGAACTTGTGTTGGAAGCA 1168
QY 1006 AACTATGCTGTGAGGAGATGTCAGGACCAATGGAATTTGGCAAGAAACAATCCCTGG 1065
DB 1169 AAATACGGTATAGAGGAATGTCAGGGAATTTGTGGAATTTGGCAAGAAACAATCCATGG 1228
QY 1066 GGTTCACGCGGGAAGCCCTTGAATTAGCTTGAAGAAATATCTGCTCATCTACGAATGCTTT 1125
DB 1229 GGATACAGGGAAGCAATTAATGAATTTGGTTGTACTTTGGAGGAGTCTTCTCCACC 1288
QY 1126 TCACATAAATAGGATCTCTTATATAGAACTTTTATGTTTATGATGATTTAGGCAG 1185
DB 1289 ATATCATATCATCTCCACTCCAGATATATATGTTGATGTTCTTTTGTGATACAAATTG 1348
QY 1186 TTCGTTGTATAATCTTCACAAATAAAATTT 1214
DB 1349 ATTGAATGATTTTACACTACGATGTT 1377

RESULT 9
US-10-425-114-29813
; Sequence 29813, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21153131B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29813
; LENGTH: 1575
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC090E09_FLI
US-10-425-114-29813

Query Match 33.0%; Score 433; DB 12; Length 1575;
Best Local Similarity 65.5%; Pred. No. 2.4e-107;
Matches 666; Conservative 0; Mismatches 345; Indels 6; Gaps 2;

QY 56 TTCTGCTCACCGGTGTCGCGGTTTTCATTGGCACCCACACCGTCGTCAGCTTCTTCAAAG 115
DB 194 TACTGTTAAACCGGCGAGCGGTTTACATCGGCAGCACACCGTTCTTTCAGCTCTTGTCTG 253
QY 116 CTGGCTTCAGCGTTTCAATATGACATTTTCGATTAATCTCGTCATGGAAGCAGTGGACC 175
DB 254 GAGGTTTCAGAGCGCGTCGTCCTCGACAACTTCGAAAATTTCTCCGAGTTGCCATGCA 313
QY 176 GGTCCCGCAAGTGGTTGGCCCTCTGCTTCTCGAAGCTTCCAAATTCACCGAGCGGATC 235
DB 314 GAGTTCAGGAGCTCGCCGCG---GAAATTTGGGAACAACCTCTCTCTTCAAGTGGAGC 370
QY 236 TCCGAAATAGGATGACATTGGAGAACTTCTTCTCCAAAGCAACATTTGATCCGCTGATCC 295
DB 371 TACGGGACAGAGCTCTCTAGACCAAAATATTTCTTCCACAAATTCGATGCTGTCTATAC 430
QY 296 ACTTTGCTGGCTTGAAGCGTTGCTGAAAGCGTTTGGCAAGCCCGCTCGCTATTTTGTAT 355
DB 431 ATTTTCTGGACTGAAAGCAGTAGAGAAAGTGTGCAAAAACCTTTTACTATATAACA 490
QY 356 TTAATTTGTTGGCACCATCAACCTCTACGAGTTTATGCGAAAGTATATTTGCAAAAGA 415
DB 491 ACAACTTGACTGGGCAATCACTCTATTTGGAAGTCAATGCTGCCCATGATGATGCAAGAAGC 550

QY 416 TGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGGCAAGATACCGTGTGGAGG 475
Db 551 TCGTGTCTTTCATCTTCAGCAACTGTATATGTTGGCCAAAGAGGTTCCATGCACAGAAG 610
QY 476 ATTTCAGTTTACAAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTGGGAAGAAATTG 535
Db 611 AGTTCCCTCTGTGACCAATGAACCCATATGGACGAACCTAAGCTTATCATTTGAAGAAATTT 670
QY 536 CCGGAGATATTGAGAAAGCTGAACAGATGAAGATCATATTACTGAGATCTTCAATC 595
Db 671 GCCGTGATGTCACCTGTGACAGCCAGATTTGTAATAATTAATTTTGAAGTACTTCAACC 730
QY 596 CAGTTGGGCTCATGAAAGTGGCAACTCGGTGAAGATCCCAAGGGCATCCCAATAAACC 655
Db 731 CAGTTGTCACACACCCAGTGGTTATTTGGGAGGATCTCGTGGAAATCCAAACATC 790
QY 656 TCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATGAGTGAACCTCAATGTATACGTC 715
Db 791 TCATGCCATTTGTTACAGCAAGTAGCAGTTTGGCCGACGCGCTGCACCTGACAGTTTGGAA 850
QY 716 ATGATTATCAACAGAGGATGGCTCTGCGATCCGGAGCTATATCCATGTGATGACTTGG 775
Db 851 ATGATTATCAACAGTGAAGTGGCACTGGGTTTCGGGATTCATTCATGTTGTTGATTAG 910
QY 776 CAGATGGCCATATTGTCGCCCTGGAAAGCTCTTCACACGGAGAACATAGTGTACTG 835
Db 911 CAGATGGGCACATTTGCTGCAATGCTTAACTAGATGAACC---TAATATAGTGTGGAGG 967
QY 836 CTTACAACTCGGAACCTGCTGGAAACATCTGCTTGAATGTTACAGCATTTGAAA 895
Db 968 TTTATNACTGGACAGAGGAGGACATCACTTTTGGAGTGTGTAGAGCTTTGAAA 1027
QY 896 AGGCTTCTGGCAAGAAATTCAGTAAATATGTCGAAGAACCGGGAGATGCGACTG 955
Db 1028 TGGCATCTGAAAGAAATTCACCTGTGTGCTGGCGTAGACCTGCTGATGCTGAAA 1087
QY 956 AGTTTATCATCTACAGAGAGCTGAGAAAGACTTGGTGGAGGCAACATATGCTG 1015
Db 1088 TTGTTTATCATCAACAAAGAAAGCGGAAGAGCTTAATGGAAAGCAAAATATGSCA 1147
QY 1016 TGGAGGAGATGTGCGAGGACCAATGGAATTTGGCCAAAGAACATCCCTGGGGTTACG 1072
Db 1148 TTGATGAGATGTGCGGTGATCAATGGAATTTGGCTAGCAAAACCCCTTATGGCTATG 1204

RESULT 10

US-10-424-599-89167
; Sequence 89167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89167
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT3847_51526C.1
US-10-424-599-89167

Query Match 32.9%; Score 431.4; DB 12; Length 1566;
Best Local Similarity 65.4%; Pred. No. 6.6e-107;
Matches 665; Conservative 0; Mismatches 346; Indels 6; Gaps 2;

QY 56 TTCTGGTACCGGTGGTGGCGGTTTCATTTGGCACCCACACCGTCTGCTTCTCAGCTTCTCAAG 115

RESULT 11
US-10-425-114-29766
; Sequence 29766, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:

Db 185 TACTGGTAAACCGGAGCCGGTTATATGGACGCCACACCGTTCTTCACTCTTCTGCTCG 244
QY 116 CTGCTTTCAGCGTTTCAATTAATCGACAAATTCGATAACTCCGTCATGGAAGCAGTGGACC 175
Db 245 GAGGTTTCAGAGCCGCTGCTCGACAACTCGAAAATTCCTCCGAGTTGCCATCCACA 304
QY 176 GCGTCCGCCAGTGGTTGGCCCTCTGCTTCTCAGAACCTCCAAATTCACCCAGGGCGATC 235
Db 305 GAGTCAGGGAGCTGCGCGGCG---GAATTTGGGAAACAACTCTCTCTTTCACAGGTGGACC 361
QY 236 TCCGGAATAGGGATGACTTTGGGAAACTCTTCTCCAAAACAAATTTGATGCGCGTATCC 295
Db 362 TACGGACAGAGCTGCTTAGACCAATATTTCTTCCACAAATTCGATGCTGTCTATAC 421
QY 296 ACTTTCGCTGCTTGAAGCGTTGCTTGAAGCGTTGGGAGCCCGTCTGCTATTTTGAAT 355
Db 422 ATTTTGTGCTGACTGAAAGCAGTAGGAGAAAGTGTGCAAAACCTTACTATATATAACA 481
QY 356 TTAATTTGTTGGTGGCACCATCAACTCTACGAGTTTATGGCAAGATATAATTCGAAAAGA 415
Db 482 ACAACTTGACTGGACCAATCACTCTATTTGGAAGTCACTGGCTGCCATGGATGCAAGAAGC 541
QY 416 TGGTTTCTCATCTGCAACCGTTTATGGCCAACTGGAAGATACCGTGTGGAGG 475
Db 542 TCGTGTCTTCTCATCTTTCAGCAACTGTATATGTTGGCCAAAGGAGTTCCATGACAGAAG 601
QY 476 ATTTCAAGTTTACAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTCGGAAGAAATTG 535
Db 602 AGTTCCCTCTGTGACCAATGAACCCATATGGACCACTAAGCTTATCAATGAAGAAATTT 661
QY 536 CCGGAGATATTGAGAAAGCTGAACAGATGGAAGATCATATTACTGAGATCTTCAATC 595
Db 662 GCCGTGATGTCCTGTCAGAGCCAGATTTGTAATAATTAATTTTGTAAAGATATCTCAACC 721
QY 596 CAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCAAAATAACC 655
Db 722 CAGTTGGTGACACCCCGAGTGGTTATTTGGGGAGGATCTCTGCGAATTCCAAACAATC 781
QY 656 TCATGCTTACATTCAGCAAGTAGCTGTGGAAAGATGAGTGAACCTCAATGTATACGGTC 715
Db 782 TCATGCCATTTGTTTCAGCAAGTAGCAGTTTGGCCGACGGCTGCACTGACAGTTTGGAA 841
QY 716 ATGATTATCAACAGGAGTGGCTGCGATCCGGACTATATCCATGTGATGAGACTTGG 775
Db 842 ATGATTATTAACAAGTATGGCACTGGGGTTCCGGATTCATTCATGTTGTTGATTAG 901
QY 776 CAGATGGCCATATTGTCGCCCTGCGAAAGCTCTTCAACGGAGAACATAGTGTGACTG 835
Db 902 CAGATGGGCACATTTGCTGCAATTTGCTTAAACTAGATGAACC---TAATATAGTGTGAGG 958
QY 836 CTTTACAACTGGGAACTGCTGCTGGAACATCTGCTTGAATGTTTACAGCAATTTGAAA 895
Db 959 TTTATTAACCTGGGAACAGGAAGGGAACATCAGTTTGGAGATGTTAGAGCTTTTGAAA 1018
QY 896 AGGCTTCTGGCAAGAAATTTCCAGTAAATTTATGTCCAAAGAACCGGAGATGCGACTG 955
Db 1019 TGGCATCTGGAAAGAAATTTCCACTTGTGATGGCTGGCCGTAGACCTGCTGATGCTGAAA 1078
QY 956 AGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTTGGAGGCAAACTATGGTG 1015
Db 1079 TTGTTTATGCACTCAACAAAGAAAGCGGAAGAGCTTAAATGGAAAGGCAAAATATGGCA 1138
QY 1016 TGGAGGAGATGTGACGGGACCAATGGAATTTGGCCAAAGAACCAATCCCTGGGTTACG 1072
Db 1139 TTGATGAGATGTGCGCTGTGATCAATGGAATTTGGCTAGCAAAACCCCTTATGGCTATG 1195

APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 29766
LENGTH: 1622
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GWNOIR019B10_FLI
US-10-425-114-29766

Query Match 32.9%; Score 431.4; DB 12; Length 1622;
Best Local Similarity 65.4%; Pred. No. 6.8e-107;
Matches 665; Conservative 0; Mismatches 346; Indels 6; Gaps 2;

QY 56 TTCTGTCACCGGTGGTGGCGTTTCATTGGGACCCACACCGTCGTTTCAGCTTCTCAAG 115
DB 195 TACTGGTAACCGGAGCGGTTACATCGGCAGCCACACCGTCTTCAGCTTCTTCGTCG 254
QY 116 CTGGCTTCAGCGTTCAATTAATCGACAATTCGATACTCCGTCATGGAAGCAGTGGACC 175
DB 255 GAGGTTTCAGAGCGGTCGCTTCGACACCTCGAATTCCTCCGAGGTTGCCATCCACA 314
QY 176 GCGTCGCGCAAGTGGTGGCCCTCTGCTTCTCAGAACTCCAACTCCACCCAGGGGATC 235
DB 315 GAGTCAGGAGCTCGCGGC--GAATTTGGGAACAACCTCTCTTTCACAAAGGTGGACC 371
QY 236 TCCGGAATAGGATGACTTCGGGAACCTCTCTCCAAACAACATTTGATGCGGTGATCC 295
DB 372 TAGGGACAGAGTGTCTAGACCAATATTTCTTCCACAAATTCGATGCTGTCATAC 431
QY 296 ACTTGTGCTGTAAGAGCGTTGCTGAAAGCGTTGCGAAGCGCCGTCGCTATTGTGAT 355
DB 432 ATTTGTGCTGACTGAAAGCAGTAGAGAGAGTGTGCAAAACCTTACTATATACTATAACA 491
QY 356 TTAATTTGGTGGCACCATCACTCTACGAGTTATGTCMAAGTATATTCGAAAGA 415
DB 492 ACACTTGACTGGGACAATCACTCTATGGAAGTCAATGCTGCCATGATGCAAGAAGC 551
QY 416 TGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTGAAAGATACCGTGTGAGGAGG 475
DB 552 TCGTGTTCATCTTCAGCACTGTATATGTTGGCCAAAGGAGGTTCCATGCACAGAG 611
QY 476 ATTTCAGTTACAGCTATGAATCCCTATGGAGCGGACCAAGCTTTTCCTGGAGAAATG 535
DB 612 AGTTCCCTCTGTGAGCAATGAACCCATATGGACGAACTAAGCTTATCATTTGAAGAAAT 671
QY 536 CCGAGATATTGAAAGCTGAAACGAGATGGAAGATCATATTACTGATATCTTCAATC 595
DB 672 GCGGTGATGTCCTGTGAGAGCCAGATTTGAAATATTTTGTAAATATCTTCAACC 731
QY 596 CAGTGGGGCTCATGAAAGTGGCAACTCGGTGAAGATCCCAAGGCGATCCCAATAAAC 655
DB 732 CAGTTGGTGACACCCCGAGTGGTTATATTTGGGGAGGATCTCGTGAATTCGAAACAATC 791
QY 656 TCATGCTTACATTACAGCAAGTAGCTGTTGGAAGATTGACTGAATCAATGATATACGGTC 715
DB 792 TCATGCCATTTGTTACGCAAGTAGAGTGGCCGACGGCTGCATGACAGTTTGTGNA 851
QY 716 ATGATTATCAACAGAGGATGCTCTGCGATCCGGAGCTATATCCATGATGATGGACTGG 775
DB 852 ATGATTATATACAGATGATGGCACTGGGGTTCGGGATTTACATTCATGTTGATTTAG 911
QY 776 CAGATGGCCATATTGCTGCCCTCGGAAGCTCTTTCACACGGGAGACATAGGTTTACTG 835

DB 912 CAGATGGGCACATTTGCTGCTTAACTAGATGAACC--TAATATAGTTGTGAGG 968
QY 836 CTTACAACTGGGAACCTGGTGGTGAACAATCTGTGCTTGAATGTTTACAGCATTTGAAA 895
DB 969 TTTATAACCTGGGAACAGGAAGGAAACATCAGTTTTTGGAGATGTTTAGAGCTTTTGA 1028
QY 896 AGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCCAAAGAACCGGAGATCGGACTG 955
DB 1029 TGGCATCTGGAAGAAATTCACCTTGTGATGGCTGGCCGTAGACCTGTTGATGCTG 1088
QY 956 AGGTTTATGCTATACAGAGAGACTGAGAAAGAACTTTGTTGGAAGGCAAACTATGTTG 1015
DB 1089 TTGTTTATGATCAACAAAGAAAGCGGAAGAGAGCTTTAAATGGAAGCAAAATATGCA 1148
QY 1016 TGGAGGAGATGTCAGGAGCAATGGAATGGGCAAGAAACAATCCCTGGGGTTACG 1072
DB 1149 TTGATGAGATGTGCGGTGATCATGGAATGGGCTAGCAAAACCCCTTATGGCTATG 1205

RESULT 12
US-10-425-114-5152
; Sequence 5152, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 5152
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700445105_FLI
US-10-425-114-5152

Query Match 32.8%; Score 430.2; DB 12; Length 1497;
Best Local Similarity 65.0%; Pred. No. 1.4e-106;
Matches 668; Conservative 0; Mismatches 353; Indels 6; Gaps 2;

QY 54 CATTCTGTCTACCGGTGGTGGCGTTTCATTGGCACCACACCGTCGTTTCAGCTTCTCAA 113
DB 171 CATCTCGTGACGGCGGCGCGGTTACATCGGCAGCCACACCGTGTGCTGCTGCTGCA 230
QY 114 AGCTGGCTTCAGCGTTTCATATTCGACAATTCGATACTCCGTCATGGAAGCAGTGA 173
DB 231 GCAGGGCTTCGCGCTCGTCTGTCGACAACCTCGACAACGCTCCGAGGCGCCCTCGC 290
QY 174 CCGGTCGCGCAAGTGGTGGCCCTCTGCTTCTCAGAACTCCCAATTCACCCAGGGCGA 233
DB 291 CCGGTGCGCAAGTGGTGGCCCTCGCGGCGACAGCGGC--CAACTGCTTCCACAGTTGA 347
QY 234 TCTCCGGAATAGGATGACTTGGAGAACTTCTTCTCCAAACCAACATTTGATGCGGTGAT 293
DB 348 CTTTCGCGACAGGACCGCTTGGTGGACATCTTCTCGTCGACAGGTTGAGGCTGTGAT 407
QY 294 CCATCTTCTGGCTTGAAGCGGTTGCTGAAAGGTTTGGAGCCCGCTCGCTATTTTGA 353
DB 408 TCATTTCTGGGCTCAAGGCTGTTGGGAGAGCGGTGCAACAGCCCTCTACTTTACTAGA 467
QY 354 TTTTAAATTTGGTTGGCAACCACTCTACGAGTTTATGCAAGATATTAATTTGCAAAA 413
DB 468 CAACAACCTGCTGGCCACCATCACCTCTCTGGAGGTGATGCTGCGAACCGCTGCAAGAA 527

414 GATGGTTTCTCATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGA 473
Db
528 GCTGGTGTCTGCTCATCTGCAACTGTCTATGGTGGCCCAAGGAGTACCTTGACCGGA 587
Qy 474 GGATTTCAAGTTACAGCTATGAATCCCTATGGAAGGACCAAGCTTTTCCCTGGAGAAAT 533
Db 588 AGAATTTCCGCGCTCTGCGCCACCAATCCCTATGCGGCGGACAAAGCTTGTGATTTGAAGACAT 647
Qy 534 TGCCCGAGATATTTCAGAAAGCTGAAACAGAAATGGAAGATCATATTACTGAGATCTTCAA 593
Db 648 CTGCCCGAGCTGCCACCGCTCGACCCCGACTGGAGATCATCTGCTCAGGTACTTCA 707
Qy 594 TCCAGTTGGGGTCTATGAAGTGGCAACTCGGTGAAGATCCCAAGGGCATCCCCAAATAA 653
Db 708 CCGGTTGGCGTCTATCCAAAGCGGTACATCGCGAAGACCCCTGCGGTGTCCCGAACAA 767
Qy 654 CCTCATGCTTACATTCAGCAAGTAGCTGTGGAAGATTGACTGAATCAATGTATACGG 713
Db 768 CTTGATGCTTACGTGCGACAAAGTCTGCTTGGAGGTTTACCTCACGCTCAGGTCTACGG 827
Qy 714 TCATGATTATCCAAACGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
Db 828 GACGGACTACAGCACCAAGATGGGACTGGGTGGGTGATTTATATCCACGTTGCCGACT 887
Qy 774 GCGAGATGCCATATGCTGCCCTGCGAAGCTCTTCA---ACGGAGACATAGTTG 830
Db 888 GGTGACGGCCACATAGAGCCCTGAGGAAGCTCTACGAAGACTCCGACAAATAGGCTG 947
Qy 831 TACTGTCTTCAACCTGGGAACCTGGTCTGGAACATCTGTGCTTGAATGGTTTACAGCAT 890
Db 948 TGAAGTGTACAACTTGGGACTGGAAAGGGGAGCTCGGTGTTGGAAATGGTGTGCATT 1007
Qy 891 TGAAGAGCTTCTGGCAAGAAATCCAGTAAATATGTCTCAAGAGACCGGAGATGC 950
Db 1008 CGAAGAGGCTTCTGGGAAGAAATCCCTCTGTGTGCTGGCGCAAGACCGGAGAGCG 1067
Qy 951 GACTGAGGTTTATGCACTACAGAGAGCTGAGAAAGAACTTGGTTGGAAGGCAAACTA 1010
Db 1068 AGAGATCGTCTACGCGCAACTCCGACGACAGAGAGAGCTCAATGGAAGGCCAAGTA 1127
Qy 1011 TGGTGTGGAGGAGATGTGAGGACCAATGGAATTTGGGCAAGACATCCCTGGGGTTA 1070
Db 1128 CGGGATCGAGGAGATGTGAGAGATCTGTGGAATCTGGGAGAGACCGGTACGGGTA 1187
Qy 1071 CGCGGGG 1077
Db 1188 CGCTGGG 1194

RESULT 13

US-10-425-114-1181
; Sequence 1181, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1181
; LENGTH: 1531
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI
US-10-425-114-1181

Query Match 32.8%; Score 430.2; DB 12; Length 1531;
Best Local Similarity 65.0%; Pred. No. 1.4e-106; Indels 6; Gaps 2;
Matches 668; Conservative 0; Mismatches 353;

Qy 54 CATCTGGTTCACCGGTGGTCCCGTTTCATTTGGCAACCCACACCGTCTGCTTCAAGTCTTCAA 113
Db 177 CATCTCGTACCGGGGGCGCGGTACATCGGCAGCCACACCGTCTGCGAGCTGTGCA 236
Qy 114 AGCTGGCTTCAGAGTTTCAATATTCAGAAATTTTCGATAATCTCGTTCATGGAACAGTGA 173
Db 237 GCAGGGCTTCGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 236
Qy 174 CCGCGTCCGCAAGTGGTCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 233
Db 297 CCGCGTCCGCGAGCTGCGCGGCGACGAGCGCG---CAACCTGCTCTTCCACAAGGTGA 353
Qy 234 TCTCCGGAATAGGATGACTTTGGAGAACTCTTCTCCAAACAACTTGTGATGCGCTGAT 293
Db 354 CTTTCGCGACAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 413
Qy 294 CCACTTTGCTGGCTTGAAGCGTTGCTGAAAGCGTTGGAAGCGTGGAAAGCGCGCTGCTG 353
Db 414 TCACTTTGCTGGCTCAGGCTGCTGCGGAGAGCGTGCAACAGCCCTACTTTTACTACGA 473
Qy 354 TTTTAATTTGGTGGCAACCATCACTCTACAGAGTTTATGGCAAGTATATATTTGCAAAA 413
Db 474 CAACAACCTGCTCGGACCATCACTCCCTCTGAGAGTGTGGCTGCGAAACGGCTGCAAGAA 533
Qy 414 GATGGTTTCTCATCTGCAACCGTTTATGCGCAACCTGAAAGAGATACCGTGTGAGGA 473
Db 534 GCTGGTCTCTCTGCTCATCTGCAACTGTCTATGGTGGCCCAAGAGTACCTGCTGACGTA 593
Qy 474 GGATTTCAAGTTACAGCTATGAATCCCTATGGAAGCGACCAAGCTTTTCTGGAAGAT 533
Db 594 AGAATTTCCGCTCTGCGGACCAATCCCTATGCGGCGGACAAAGCTTTGTGATTTGAAGAC 653
Qy 534 TGCCCGAGATATTTCAGAAAGCTGAAACCGTTTATGCGCAACCTGAAAGAGATACCGTGTG 593
Db 654 CTGCGCGAGCTGCAACCGCTCCGACCCGACTGGAAGATCATATCTGCTCAGGTACTTCAA 713
Qy 594 TCCAGTTGGGCTCATGAAAGTGGCAAACTCGGTGAGATCCCAAGGCAATCCCAATAA 653
Db 714 CCGGTTGGGCTCATGCAAGCGGTACATCGCGAAGACCCCTGCGGTGTCCCGAACAA 773
Qy 654 CCTCATGCTTACATTCAGCAAGTGTGTTGGAAGATTGACTGAACTCAATGTATACGG 713
Db 774 CTTGATGCCCTACGTGCGAGCAAGTGTGTTGGAAGTTACCTCACTCAGCGTCTACGG 833
Qy 714 TCATGATTATCCAAACGAGGATGGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
Db 834 GACGGACTACAGCAACCAAGATGGGACTGGGTGCTGATTTATCATCCACGTTGCGACCT 893
Qy 774 GCGAGATGGCCATATTGCTGCCCTGCGAAGCTCTTCA---ACGGAGAACATAGTTG 830
Db 894 GGTGACGGCCCATAGAGCCCTGAGGAAGCTCTAGGAAGACTTCAGGAAGACTCCGACAAATAG 953
Qy 831 TACTGCTTACAACTCGGAACTCGTCTGGAACATCTGTGCTTTGAAATGGTTACAGATT 890
Db 954 TGAAGTGTACAACTTGGGACTGGAAGGGGAGCTCGGTGTTGGAATGGTGTGTTATT 1013
Qy 891 TGAAGGCTTCTGCGAAGAAATTCAGTAAATATGTCCAAAGAGACCGGGAGATGC 950
Db 1014 CGAAGGCTTCTGGGAAGAAATTCCTCTGTTGCTGGCGGAGAGACCCGAGAGCG 1073
Qy 951 GACTGAGGTTTATGCACTTACAGAGAGCTGGAAGAACTTGGTTGGAAGCAAACTA 1010
Db 1074 AGAGATGCTCTACGCCCACTCCCAAGGCAGAGAAGAGCTCAATGGAAGGCCAAGTA 1133
Qy 1011 TGGTGTGAGGAGATGTGAGGACCAATGGAATTTGGGCAAGAACTTCCCTGGGGTTA 1070
Db 1134 CGGATCGAGGAGATGTGAGAGATCTGTGGAATCTGGCGAGCAAGACCCCTGACGGTA 1193

QY 1071 CGCGGG 1077
DB 1194 CGTGGG 1200

RESULT 14
US-10-425-114-22869
; Sequence 22869, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22869
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-035-C3_FLI
US-10-425-114-22869

Query Match 30.8%; Score 403.8; DB 12; Length 1200;
Best Local Similarity 64.8%; Pred. No. 1.9e-99;
Matches 616; Conservative 0; Mismatches 332; Indels 3; Gaps 1;

QY 54 CATTCTGGTCCCGGTGGTGGCGGTTTCATTGGCACCCACACCGTGGTTTCAGCTTCTCAA 113
DB 253 CGTGTGGTCAAGGGGGGGGGTACATCGTAGCCACGGCGTGTCTGAGCTCTCTCT 312
QY 114 AGCTGGCTTCAGGTTTCAATATCCAAATTCGATRACTCGGTGATGAGGAGTGA 173
DB 313 CGGGGGCTCCCGGGCGGTGTCATGACCACTCAACAACTCTCGAGCTGGCGTCCG 372
QY 174 CCCTGTCGCGCAAGTGGTGGCGCTCTGCTTTCTCAGAACCTTCAATCCACCGAGGGA 233
DB 373 CGCGTGGC---GCGCTCGGGGGACCACTCCCGCACTCTCTTCCACAAGATTGA 429
QY 234 TCTCCGAATAGGATGACTTGGAGAACTTCTCCAAAACAACATTTGATGCGGTGAT 293
DB 430 TCTCCGTGACAGGGGACCTGGAAATGGTTTGTCTTCTACAAGATTTGAAGCTGTAT 489
QY 294 CCACTTTGCTGGCTTGAAGCGGTTCTGAAAGCGTTGCGAAGCCCGTGGCTATTTTGA 353
DB 490 TCACCTTCGTGGATTGAAGCTGTGGTGAAGCGTACAGAGCCATTACTTTATTATGA 549
QY 354 TTTTAATTTGGTGGCACCATCACTCTACAGATTTATGGCAAGTATATTTGCAAAA 413
DB 550 CAACAACGTCATTGGCACGATTAATCTCTAGAAGTTATGTCTGTTCACGGTTGAAGAA 609
QY 414 GATGGTTTCTCATCTGCAACCGTTTATGGCAACCTGAAAGATACCGTGTGAGGA 473
DB 610 GTTGGTGTCTCATCACTGCTGAGTTTATGGATCACCCAAAACCTACCTCTGCACGA 669
QY 474 GGAATTTCAAGTTACAGCTATGAATCCCTATGGACGACCAAGCTTTTCTGGAGAAT 533
DB 670 AAATTTTCTTACTTCAACAATCCATATGGCAAAACAAAGCTGTGTTTGAAGATAT 729
QY 534 TGCCTCAGATATTCAGAAAGCTGAACCAAGATGGAAGATCATATTACTGAGATCTTCAA 593
DB 730 TTCCCGGATATCTACCGTTTCAATCTGAAATGGAAGATCATTTTACTTAGGTACTTCA 789
QY 594 TCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGGATCCCAATAA 653
DB 790 TCCAGTTGGTGTCTCATCTAGTGAATCTTGGCGAGGACCCACGAGGAATTTCCCAACA 849

QY 654 CCTCATGCTTACATTGACCAAGTAGCTGTTGGAAGATTGACTGAACCTCAATGTATACGG 713
DB 850 TCTTATGCCCTATGTTTACGCAAGTTGGGTTGGTAGAGGCCAGCTCTAACAGTTTAGG 909
QY 714 TCATGATTATCCAAAGAGGATGCTCTGCGATCCGGGACTATATCCATGTGATGACTT 773
DB 910 AATGACTATGCAACAAGAGATGGGACTGGGTCCGAGATTACATCATGTGTTGACCT 969
QY 774 GGCAGATGGCCATATTGCTGCCCTGCGAAAGCTCTTCCAAACGAGAACATAGTTGTAC 833
DB 970 TGTGACGGACATATTGCTGCAATTGCAAGAGCTTTTGAAGCTCTAGCATAGGTTGA 1029
QY 834 TGCTTACAACCTGGGAACTGGTGGTGGAAATCTGTGCTTGTGAAATGTTTACAGCATTTGA 893
DB 1030 AGCTGTAACACTTGGAAACCGGAGAGGTACATCTGTCTGCGAGATTCTTAAGCATTTGA 1089
QY 894 AAGGCTTCTGGCAAGAAATTCAGTAAATATATGTTCCAAAGAACCGGAGATGCGAC 953
DB 1090 GAGGCTTCTGGGAAGAAATATCTCTGATTTTGGTGAAGAGCCCGAGTGTATGAGA 1149
QY 954 TGAGGTTTATGCACTCTACAGAGAGCTGAGAAGAACTTTGGTTGGAAGC 1004
DB 1150 GATTCTGTTTTCAGAGACTACTAAAGCAGAGAGGAGCTTAACCTGGAAGC 1200

RESULT 15
US-10-425-114-3333
; Sequence 3333, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3333
; LENGTH: 1729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700260760_FLI
US-10-425-114-3333

Query Match 30.7%; Score 402.2; DB 12; Length 1729;
Best Local Similarity 65.3%; Pred. No. 6.7e-99;
Matches 590; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 170 TGAACCGGTCCCGCAAGTGGTGGCGCTCTGCTTCTCAGAACCTTCAATTCACCCAGG 229
DB 133 TGCAGCTCTCTCTCGCGGGTTCGCGCGCTCTGCTATCGAACCTCTCTTCCACAAGA 192
QY 230 GCGATCTCCGGAATAGGAGTGAATTTGGAGAACTCTTCTCCAAAACAACATTTGATCCG 289
DB 193 TTGATCTCCGTGACAGGGAGCACTGGAATGGTTTGTCTTCTACAGATTGAAGCTG 252
QY 290 TGATCCATTTGCTGGCTTGAAGCGGTGCTGAAAGCGTTGCGAACCGCGTGCCTATT 349
DB 253 TCATTCATCTCGCTGGATTGAAAGCTGTGGGTGAAAGCTGACAGAACCTTACTTTATT 312
QY 350 TTGATTTTAAATTTGGTGGCACCATCAACCTCTACAGATTTATGGCAAGATATAATTGCA 409
DB 313 ATGCAACAACGTCATTGGCACGATTAATCTTCTAGAGTTATGTCTGTTCACGGTTGCA 372
QY 410 AAAAGATGGTTTCTCATCATCTGCAACCGTTTATGGCCAACTTGAAAAGATACCGTGTG 469

```

Db      373 AGAAGTGGTGTCTCATCATCAGCTGCAGTGTATGGATCAACCAAACTCACCCCTGCA 432
QY      470 AGGAGGATTTCAAGTTTACAAGCTATGAATCCCTATGGACGGACCAAGCTTTTCTGGGAAG 529
Db      433 CAGAAATTTTCTCTTACTCCAAACATCCATATGGCAAAACAAGCTCGTTGTTGAAG 492
QY      530 AAATGCCCGAGATATTCAAGAGCTGAACAGAGATGGAAGATCATATTAATCTAGATACT 589
Db      493 ATATTTGCCGGATATCTACCGTTCCAGATCCTGAATGGAAGATCAATTTTACTTAGGTACT 552
QY      590 TCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCAAGGGCATCCCAA 649
Db      553 TCATCCAGTTGGTGTCTCATCTAGTGGATATCTTGGCGAGGACCCACGAGGAATTCOCA 612
QY      650 ATAACTCATGCTTACATTCAGCAAGTAGCTGTTGGAGATGTAAGTGAATCTCAATGTAT 709
Db      613 ACAATCTTATGCCCTATGTTTCAAGAGTTGCGGTTGGTAGGAGGCCAGCTCTAACAGTTT 672
QY      710 ACGGTCTATGATTCACCAAGGAGGTGGCTCTGCGATCCGGGACTATATCCATGTGATGG 769
Db      673 TAGGAATGACTATGCAACAGAGATGGGACTGGGTCGAGATTAATCCATGTGGTTG 732
QY      770 ACTTGGCAGATGGCCATATTGTGCCCTCGAAAGCTCTTCAACGGAGAACATAGGTT 829
Db      733 ACCTTGCTGACGGACATATTGCTGCATTGCAGAGCTTTTGAGAACTCTAGCATAGGGT 792
QY      830 GTACTGCTTACACCTGGGAACCTGGTCTGGGACATCTGTGCTTGAATGGTTACAGCAT 889
Db      793 GTGAGCGTACACCTTGGGAACCGGAAGAGGTACATCTGTGCTGGAGATTGTTAAAGCAT 852
QY      890 TTGAAAGGCTTCTGGCAAGAAATTCAGTAAATTTATGTCCAAGAGACCGGGAGATG 949
Db      853 TTGAGAAGGCTTCTGGGAAGAAATACCTCTGATTTTGGTGAAGAGCGCCAGGTGATG 912
QY      950 CGACTGAGGTTTATGCTATCTACAGAGAGCTGAGNAGAACTTGGTTGGAGGCAACT 1009
Db      913 CAGAGATTCTGTTTTCAGAGACTACTAAAGCAGAGAGGGAGCTTAACCTGGAAGCAAAAT 972
QY      1010 ATCGTGTGGAGGAGATGTGAGGAGCAATGGAATGGCAAGAACCAATCCCTGGGGTT 1069
Db      973 ACGGTATTGAGAGATGTGCGCGACCAATGGAACTGGGCCAGCAAGAACCTTATGGCT 1032
QY      1070 ACG 1072
Db      1033 ATG 1035

```

Search completed: March 22, 2004, 21:52:30
 Job time : 498 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 22, 2004, 17:30:18 ; Search time 1766 Seconds
(without alignments)
10403.394 Million cell updates/sec

Title: US-09-913-064A-13
Perfect score: 1312
Sequence: 1 gacagagccactctctccc.....aaaaaaaaaaaaaaaaaaaa 1312

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:**

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estnu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_hic:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_hic:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_eston:**
- 17: em_gss_hum:**
- 18: em_gss_inv:**
- 19: em_gss_pln:**
- 20: em_gss_vit:**
- 21: em_gss_fun:**
- 22: em_gss_mam:**
- 23: em_gss_mus:**
- 24: em_gss_pro:**
- 25: em_gss_rod:**
- 26: em_gss_phg:**
- 27: em_gss_vrl:**
- 28: gb_gss1:**
- 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	765.4	58.3	775	14	CA802698 sau40g01.
2	636.8	48.5	664	10	AW508396 s140d07.y
3	599.2	45.7	649	10	BF324112 su228c4.y
4	584.8	44.6	613	10	BF324111 su229i2.y

5	573.4	43.7	575	12	BM521748
6	571.4	43.5	574	12	BM086182
7	562.8	42.9	567	10	AW472274
8	561.4	42.8	564	10	AW569126
9	560.2	42.7	798	14	CB892692
10	558.8	42.6	580	10	AW507982
11	558.4	42.6	564	9	A1856802
12	550.8	42.0	560	10	BE474654
13	540.6	40.1	547	12	BM528276
14	535.2	41.8	694	10	BF635019
15	532	40.5	552	12	BM891597
16	531	40.5	718	12	BG644768
17	530.4	40.4	879	13	BQ797360
18	525.6	40.1	718	12	BG647665
19	519.6	39.6	685	12	BG455279
20	518.4	39.5	564	12	B1498305
21	510	38.9	555	10	BE352750
22	509.6	38.8	551	12	B1946267
23	502	38.3	690	10	AW775957
24	501	38.2	613	12	BM813737
25	498.2	38.0	648	12	B1271421
26	495.2	37.7	626	12	B1270061
27	491.4	37.5	493	12	B1944852
28	491.2	37.4	560	12	BM177819
29	490	37.3	526	12	B1942090
30	489.6	37.3	721	14	CF209053
31	484.4	36.9	812	12	BG584474
32	481.4	36.7	651	10	AW774538
33	479.2	36.5	509	10	BE661026
34	478.2	36.4	753	13	BU635732
35	468.6	35.7	579	14	CB894137
36	465.2	35.5	672	10	BF645662
37	454.2	34.6	475	10	AW507725
38	451.8	34.4	475	10	AW507725
39	450	34.3	566	12	B1270893
40	447.2	34.1	719	14	CD841689
41	443	33.8	1816	11	AY104881
42	434	33.1	899	14	CK284851
43	433.4	33.0	727	13	BU018048
44	431	32.9	431	12	B1943653
45	429.6	32.7	437	10	AW397514

ALIGNMENTS

RESULT 1	CA802698	775 bp	mRNA	linear	EST 05-DEC-2002
CA802698	sau40g01.y1	Gm-cl071	Glycine max	cDNA clone	SOYBEAN CLONE ID:
LOCUS	Gm-cl071-3386	5'	similar to	TR:Q43070	Q43070
DEFINITION	UDP-GALACTOSE-4-EPIMERASE ;		mRNA sequence.		
ACCESSION	CA802698				
VERSION	CA802698.1				GI:26059784
KEYWORDS	EST.				
SOURCE	Glycine max (soybean)				
ORGANISM	Glycine max				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
AUTHORS	1 (bases 1 to 775)				
	Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
TITLE	Public Soybean EST Project				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine				

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estowatson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 456.
 Location/Qualifiers

FEATURES

1..775
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl071-3386"
 /issue type="Immature pods (~2cm long) of greenhouse
 grown plants"
 /lab_host="DH10B"
 /clone_lib="Gm-cl071"
 /notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; The
 cDNA library was constructed from mRNA isolated from
 immature pods (approximately 2cm long) of greenhouse grown
 plants. The library was prepared using the Life
 Technologies pSuperScript cDNA library construction kit.
 Complementary DNA was synthesized from mRNA using a
 poly(dT) sequence with a NotI restriction site. SalI
 linkers adapters were ligated to the blunt-ended cDNA
 fragments followed by NotI digestion. The cDNA fragments
 were directionally cloned into the NotI-SalI restriction
 site of the pSPORT1 vector. The ligated cDNA fragments
 were transformed into E.coli Electromax DH10B host cells.
 This library was constructed in the laboratory of Dr. Lila
 Vodkin by Anu Khanna at the University of Illinois at
 Urbana-Champaign. email: l-vodkin@uiuc.edu"

ORIGIN

Query Match 58.3%; Score 765.4; DB 14; Length 775;
 Best Local Similarity 99.2%; Pred. No. 2.8e-121;
 Matches 769; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 39 GTCTTCCTCCACACATCTGCTACCGGTGGTGGCGGTTTCATTGGACCCACACCGT 98
 Db 1 GTCTTCCTCCACACATCTGCTACCGGTGGTGGCGGTTTCATTGGACCCACACCGT 60

QY 99 CGTTACGCTCTCAAGCTGGCTTCAGCGTTTCAATAATCGAAATTCGATAATCCGT 158
 Db 61 CGTTACGCTCTCAAGCTGGCTTCAGCGTTTCAATAATCGAAATTCGATAATCCGT 120

QY 159 CATGGAAGCAGTGGACCGGTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCA 218
 Db 121 CATGGAAGCAGTGGACCGGTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCA 180

QY 219 ATTCACCCAGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACAC 278
 Db 181 ATTCACCCAGCGGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAACAC 240

QY 279 ATTGTATGCGGTGATCCATTTGCTGGCTTGAAGCGTTGCTGAAAGCGTTGCGAGCC 338
 Db 241 ATTGTATGCGGTGATCCATTTGCTGGCTTGAAGCGTTGCTGAAAGCGTTGCGAGCC 300

QY 339 CCGTCGCTATTTGATTTTAAATTTGGTTGGCACCATCACTCTACGATTTATGCAAA 398
 Db 301 CCGTCGCTATTTGATTTTAAATTTGGTTGGCACCATCACTCTACGATTTATGCAAA 360

QY 399 GTATAATTGCAAAAGATGTTTTCTCATCATCTGCAACCGTTTATGGCAACCTGAAA 458
 Db 361 GTATAATTGCAAAAGATGTTTTCTCATCATCTGCAACCGTTTATGGCAACCTGAAA 420

QY 459 GATACCGTGTGAGAGATTTCAAGTTACAGTATGATCCCTATGAGCGGACCAAGCT 518
 Db 421 GATACCGTGTGAGAGATTTCAAGTTACAGTATGATCCCTATGAGCGGACCAAGCT 480

QY 519 TTTCTCTGGAAGAAATTGCCCGAGATATTTCAGAAAGCTGAACCAAGTGAAGATCATATT 578
 Db 481 TTTCTCTGGAAGAAATTGCCCGAGATATTTCAGAAAGCTGAACCAAGTGAAGATCATATT 540

QY 579 ACTGAGATATCTTCAATCCAGTTCGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCA 638
 Db 541 ACTGAGATATCTTCAATCCAGTTCGGGCTCATGAAAGTGGCAAACTCGGTGAAGATCCCA 600

QY 639 GGGCATCCCAAAATACCTCATGCTTACATTCAGCAAGTAGCTGTGGAGATTGACTGA 698
 Db 601 GGGCATCCCAAAATACCTCATGCTTACATTCAGCAAGTAGCTGTGGAGATTGACTGA 660

QY 699 ACTCAATGTATACGGTTCATGATTATCCACAGGAGTGGCTCTGCGATCCGGGACTATAT 758
 Db 661 ACTCCATGTATACGGTTCATGATTATCCACAGGAGTGGCTCTGCGATCCGGGACTATAT 720

QY 759 CCATGTGATGAGCTTGGCAGATGGCCATATTGCTGCCCTCGCAAGCTTCTTACA 813
 Db 721 CCATGTGATGAGCTTGGCAGATGGCCATATTGCTGCCCTCGCAAGCTTCTTACA 775

RESULT 2

AW508396

LOCUS

DEFINITION

664 bp mRNA linear EST 03-DEC-2001

Gm-r1030-1382 5' similar to TR:Q43070

UNP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

AW508396

AW508396.1

GI:7146474

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

REFERENCE

1 (bases 1 to 664)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, I., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estowatson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Insert Length: 1309 Std Error: 0.00

High quality sequence stop: 441.

Location/Qualifiers

1..664

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-1382"

/lab_host="DH10B"

/clone_lib="Gm-r1030"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This

cDNA library was constructed from mRNA isolated from

immature cotyledons of greenhouse grown plants

(individual seed fresh weight of 100-300mg). The library

was prepared using the Life Technologies pSuperScript cDNA

library construction kit. Complementary DNA was

synthesized from mRNA using a poly(dT) sequence with a

NotI restriction site. SalI linkers adapters were ligated

to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPOR1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN		Query Match	48.5%; Score 636.8; DB 10; Length 664;
		Best Local Similarity	97.4%; Pred. No. 3e-99;
		Matches	647; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy	30	CAGATGGTGTCTTCCTCCCAACACATCTGGTACCGGTGGTGGCGTTTCATTTGGCAC	89
Db	1	CAGATGGTGTCTTCCTCCCAACACATCTGGTACCGGTGGTGGCGTTTCATTTGGCAC	60
Qy	90	CCACACCGTGTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCATATCGACAAATTCGA	149
Db	61	CCACACCGTGTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCATATCGACAAATTCGA	120
Qy	150	TAACTCCGTGATGAAGCAGTGGACCGGTCCGCAAGTGGTGGCCCTCTGCTTTCTCA	209
Db	121	TAACTCCGTGATGAAGCAGTGGACCGGTCCGCAAGTGGTGGCCCTCTGCTTTCTCA	180
Qy	210	GAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGAGAAATCTTCTC	269
Db	181	GAACCTCCAAATTCACCCAGGCGATCTCCGGAATAGGATGACTTGAGAAATCTTCTC	240
Qy	270	CAAAACAACATTTGATGCCGTGATCCACTTTTGTGGCTTGAAAGCGTGTCTGAAAGCGT	329
Db	241	CAAAACAACATTTGATGCCGTGATCCACTTTTGTGGCTTGAAAGCGTGTCTGAAAGCGT	300
Qy	330	TGCGAAGCCCGTGCCTATTTGATTTTAAATTTGGTTGGCCACCATCAACCTCTACGATT	389
Db	301	TGCGAAGCCCGTGCCTATTTGATTTTAAATTTGGTTGGCCACCATCAACCTCTACGATT	360
Qy	390	TATGCGCAAGTATATATGCAAAAGATGGTTTCTCATCTATCTGCAACCGTTTATGCGCA	449
Db	361	TATGCGCAAGTATATATGCAAAAGATGGTTTCTCATCTATCTGCAACCGTTTATGCGCA	420
Qy	450	ACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTCAAGCTATGATCCCTATGGACG	509
Db	421	ACCTGAAAGATACCGTGTGAGAGGATTTCAAGTTCAAGCTATGATCCCTATGGACG	480
Qy	510	GACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAACCAAGTGGAA	569
Db	481	GACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAACCAAGTGGAA	540
Qy	570	GATCATATTACTGAGATCTTCAATCCAGTTGGGGCTTCATGAAAGTGGCAATTCGGTGA	629
Db	541	GATCATATTACTGAGATCTTCAATCCAGTTGGGGCTTCATGAAAGTGGCAATTCGGTGA	600
Qy	630	AGATCCCAAGGGCATCCCAATACCTTCATGCCCTTACATTCAGCAAGTAGCTGTGGAG	689
Db	601	AGATCCCAAGGGCATCCCAATACCTTCATGCCCTTACATTCAGCAAGTAGCTGTGGAG	660
Qy	690	ATTG 693	
Db	661	AATG 664	

RESULT 3
BF324112
LOCUS
DEFINITION
su22h04.v1 Gm-cl068 Glycine max cDNA clone GENE EST 06-DEC-2001
Gm-cl068-343 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.
ACCESSION
BF324112
VERSION
BF324112.1 GI:11273736
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 649)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cduresgen.com
High quality sequence stop: 420.

Location/Qualifiers
1..649
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl068-343"
/tissue_type="Leaf, drought stressed, 1 month old plants, greenhouse grown"
/lab_host="DH10B"
/clone_lib="Gm-cl068"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN		Query Match	45.7%; Score 599.2; DB 10; Length 649;
		Best Local Similarity	98.3%; Pred. No. 8e-93;
		Matches	637; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
Qy	16	CTCCCTCTCTATTGAGCATGGTGTCTCTCCCAACACATTCGGTCAACGGTGGTGGCC	75
Db	2	CTCCCTCTCTATTGAGCATGGTGTCTCTCCCAACACATTCGGTCAACGGTGGTGGCC	61
Qy	76	GGTTTCATTGGCACCCACACCGTCTTCAAGCTGGCTTCAGGTTTCAATA	135
Db	62	GGTTTCATTGGCACCCACACCGTCTTCAAGCTGGCTTCAGGTTTCAATA	121
Qy	136	ATCGCAATTTTCGATACTCCGTCTATGAGCAGTGGACCGCTCCCAAGTGGTGGC	195
Db	122	ATCGCAATTTTCGATACTCCGTCTATGAGCAGTGGACCGCTCCCAAGTGGTGGC	181
Qy	196	CCTCTGCTTTCTCAGAACCTTCAATTCACCGAGGCGATCTCCGGAATAGGATGAC	255
Db	182	CCTCTGCTTTCTCAGAACCTTCAATTCACCGAGGCGATCTCCGGAATAGGATGAC	241
Qy	256	GAGAACTCTTCTCCAAACACATTTGATCCCGTGATCCACTTTGCTGGTGAACGG	315
Db	242	GAGAACTCTTCTCCAAACACATTTGATCCCGTGATCCACTTTGCTGGTGAACGG	301

QY 316 GTTCTGGAAGCGTTCGGAAGCCCGTCGCTATTTTGAATTTTAAATTTGGTTGGCACCATC 375
 DB 302 GTTCTGGAAGCGTTCGGAAGCCCGTCGCTATTTTGAATTTTAAATTTGGTTGGCACCATC 361
 QY 376 AACCTTACGAGTTTATGCAAGATATAATTCACAAAGATAGTTTCTCATCATCTGCA 435
 DB 362 AACCTTACGAGTTTATGCAAGATATAATTCACAAAGATAGTTTCTCATCATCTGCA 421
 QY 436 ACGTTTATGGCAACCTCAAGATACCGTGTGAGGAGTTTCAAGTTTACAAGCTATG 495
 DB 422 ACGTTTATGGCAACCTCAAGATACCGTGTGAGGAGTTTCAAGTTTACAAGCTATG 481
 QY 496 AATCCCTATGGACGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCT 555
 DB 482 AATCCCTATGGACGACCAAGCTTTTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCT 541
 QY 556 GAA-CCAGATGGAAGATCATATTAC-TCAGTACTTCAATCCAG-TTGGGGCTCATGAA 612
 DB 542 GAACCCAGATGGAAGATCATATTACATGAATTTCTCAATACAGTTTGGGGCTCATGAA 601
 QY 613 AGTGCAAACTCGTGAAGATCCCAAGGGCATCCCAATTAACCTCATG 660
 DB 602 AGTGCAAACTCGTGAAGATCCCAAGGGCATCCCAATTAACCTCATG 649

RESULT 4
 BF324111
 LOCUS su22g12.v1 Gm-c1068 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

ACCESSION BF324111
 VERSION BF324111.1 GI:11273735
 KEYWORDS Glycine max (soybean)
 SOURCE Glycine max
 ORGANISM Glycine max (soybean)

REFERENCE 1 (bases 1 to 613)
 AUTHORS Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, I., Jackson, I., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE Contact: Shoenaker R/Public Soybean EST Project
 JOURNAL Public Soybean EST Project
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 High quality sequence stop: 365.

FEATURES
 source
 1. .613
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1068-311"
 /tissue_type="Leaf, drought stressed, 1 month old plants,
 greenhouse grown"
 /lab_host="DH10B"
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; The cDNA library was constructed from mRNA isolated

from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 44.6%; Score 584.8; DB 10; Length 613;
 Best Local Similarity 97.3%; Pred. No. 2.4e-90;
 Matches 595; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 24 CTATTGAGCATGGTGTCTTCTCCCAACATCTTGGTCACCGGTGGTCCGGTTTCAT 83
 DB 1 CTATTGAGCATGGTGTCTTCTCCCAACATCTTGGTCACCGGTGGTCCGGTTTCAT 60
 QY 84 TGGCACCACACACCGTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAA 143
 DB 61 TGGCACCACACACCGTTCAGCTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAA 120
 QY 144 TTTCGATAATCTCCGTCAATGAAGCAGTGGACCCGCTCCGCAAGTGGTTGGCCCTCTGCT 203
 DB 121 TTTCGATAATCTCCGTCAATGAAGCAGTGGACCCGCTCCGCAAGTGGTTGGCCCTCTGCT 180
 QY 204 TTCTCAGAACTCTCAATTCACCCAGGGCGATCTCCGGAATAGGATGACTTTGGAGAACT 263
 DB 181 TTCTCAGAACTCTCAATTCACCCAGGGCGATCTCCGGAATAGGATGACTTTGGAGAACT 240
 QY 264 CTTCTCCAAAACAACATTTGATCCCGTATCCACTTTTGGCTTGAAGCGTTGGCTGA 323
 DB 241 CTTCTCCAAAACAACATTTGATCCCGTATCCACTTTTGGCTTGAAGCGTTGGCTGA 300
 QY 324 AAGCGTTGCCAAGCCCGTCGCTATTTGATTTTAAATTTGGTTGGCACCACCACTCTA 383
 DB 301 AAGCGTTGCCAAGCCCGTCGCTATTTGATTTTAAATTTGGTTGGCACCACCACTCTA 360
 QY 384 CGAGTTTATGGCAAGATATAATTCAAAAGATAGTTTCTCATCATCTGCAACCGTTTA 443
 DB 361 CGAATTTATGGCAAGATATAATTCAAAAGATAGTTTCTCATCATCTGCAACCGTTTA 420
 QY 444 TGSCCAACCTGAAAGATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTA 503
 DB 421 TGGCCAACTGAAAGATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTA 480
 QY 504 TGGACGGACCAAGCTTTCTCGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAACGAGA 563
 DB 481 TGGACGGACCAAGCTTTCTCGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAACGAGA 540
 QY 564 ATGAGATCATATTTACTGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACT 623
 DB 541 ATGAGATCATATTTACTGATCTTCAATCCAGTTGGGCTCATGAAAGTGGCAACT 600
 QY 624 CGGTGAAGATCC 635
 DB 601 CGATGAATCCC 612

RESULT 5
 BM521748
 LOCUS sak70f12.y1 Gm-c1036 Glycine max cDNA clone linear EST 15-FEB-2002
 DEFINITION UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.
 ACCESSION BM521748
 VERSION BM521748.1 GI:18692900
 KEYWORDS Glycine max (soybean)

ORGANISM Glycine max

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine

AUTHORS 1 (bases 1 to 575)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

Seq primer: -4ORP from Gibco

High quality sequence stop: 435.

FEATURES

source Location/Qualifiers

1..575

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-8352"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-cl036"

/notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 43.7%; Score 573.4; DB 12; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.2e-88;
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 50 AACACATTTCTGGTCACCGGTGGTGGCGTTTCATTTGGCACCCACACCGCTTCAGCTTC 109

DB 1 AACACATTTCTGGTCACCGGTGGTGGCGTTTCATTTGGCACCCACACCGCTTCAGCTTC 60

QY 110 TCAAGCTGGCTTCACGGTTTCATTAATCGACATTTTCGATACTCCGTCGGAAGCAG 169

DB 61 TCAAGCTGGCTTCACGGTTTCATTAATCGACATTTTCGATACTCCGTCGGAAGCAG 120

QY 170 TGGACCGCTCCGCCAAGTGGTGGCGCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGG 229

DB 121 TGGACCGCTCCGCCAAGTGGTGGCGCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGG 180

QY 230 GCGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAAATTTGATGCCG 289

DB 181 GCGATCTCCGGAATAGGATGACTTGGAGAACTCTTCTCCAAAACAAATTTGATGCCG 240

QY 290 TGATCCACTTTGTGGCTTGAAGCGGTGCTGCTGAAGCGTTGGGAAGCCCGCTGCTATT 349

DB 241 TGATCCACTTTGTGGCTTGAAGCGGTGCTGCTGAAGCGTTGGGAAGCCCGCTGCTATT 300

QY 350 TTGATTTTAAATTTGGTTGGCACCACCTCAACCTCTACGAGTTTATGGCAAAAGTATAATTGCA 409

DB 301 TTGATTTTAAATTTGGTTGGCACCACCTCAACCTCTACGAGTTTATGGCAAAAGTATAATTGCA 360

QY 410 AAAAGATGGTTTTCATCATCTGCAACCGTTTATGGCCCAACCTGAAAAGATACCGTGTG 469

DB 361 AAAAGATGGTTTTCATCATCTGCAACCGTTTATGGCCCAACCTGAAAAGATACCGTGTG 420

QY 470 AGGAGGATTTCAAGTTACAGCTATGAATCCCTATGACCGACCAAGCTTTCTCTGGAAG 529

DB 421 AGGAGGATTTCAAGTTACAGCTATGAATCCCTATGACCGACCAAGCTTTCTCTGGAAG 480

QY 530 AAATGTCGCCGAGATATTGAGAAAGCTGAACCAAGATGAAGATCATATTACTGAGATACT 589

DB 481 AAATGTCGCCGAGATATTGAGAAAGCTGAACCAAGATGAAGATCATATTACTGAGATACT 540

QY 590 TCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTC 624

DB 541 TCAATCCAGTTGGGGCTCATGAAAGTGGCAAACTC 575

RESULT 6

LOCUS BM086182

DEFINITION BM086182 574 bp mRNA linear EST 19-NOV-2001

DESCRIPTION sahs3602.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl036-3651 5' similar to TR:Q43070 Q43070

ACCESSION UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

VERSION BM086182

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

REFERENCE 1 (bases 1 to 574)

AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project

JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

High quality sequence stop: 434.

FEATURES

source Location/Qualifiers

1..574

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-3651"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-cl036"

/notes="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 43.6%; Score 571.4; DB 12; Length 574;
Best Local Similarity 99.7%; Pred. No. 4.8e-88;
Matches 572; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 63 CACCGTGTGCGGTTTCATTGGACCCACACCGCTGCTCAGCTTCCTCAAGCTGGCTT 122
Db 1 CACCGTGTGCGGTTTCATTGGACCCACACCGCTGCTCAGCTTCCTCAAGCTGGCTT 60

QY 123 CAGCGTTTCATTAATCGACAATTCGATACTCCGTCAATGGAAGCAGTGGACCGGTCG 182
Db 61 CAGCGTTTCATTAATCGACAATTCGATACTCCGTCAATGGAAGCAGTGGACCGGTCG 120

QY 183 CCAAGTGTGCGGCTCTGCTTCTCAGAACCTCCAAATCCACGAGGCGATCTCCGAA 242
Db 121 CCAAGTGTGCGGCTCTGCTTCTCAGAACCTCCAAATCCACGAGGCGATCTCCGAA 180

QY 243 TAGGATGACTTGGAGAACTCTTCTCCAAAACAATTTGATGCGGTGATCCACTTGC 302
Db 181 TAGGATGACTTGGAGAACTCTTCTCCAAAACAATTTGATGCGGTGATCCACTTGC 240

QY 303 TGGCTTGAAGCGGTGCTGAAGGTTGCGAGCGCCGCTGCTATTTTGAATTTAATTT 362
Db 241 TGGCTTGAAGCGGTGCTGAAGGTTGCGAGCGCCGCTGCTATTTTGAATTTAATTT 300

QY 363 GGTGGCACCATCAACCTCTACGAGTTATGCAAGATATATTCGAAAAAGATGGTTT 422
Db 301 GGTGGCACCATCAACCTCTACGAGTTATGCAAGATATATTCGAAAAAGATGGTTT 360

QY 423 CTCATCATCGAACCGTTTATGGCAACCTGAAAGATACCGTGTGAGGAGATTCAA 482
Db 361 CTCATCATCGAACCGTTTATGGCAACCTGAAAGATACCGTGTGAGGAGATTCAA 420

QY 483 GTTACAAGCTATGAATCCCTATGGACGACCAAGCTTCTCTGGAAGAAATGCCCCAGA 542
Db 421 GTTACAAGCTATGAATCCCTATGGACGACCAAGCTTCTCTGGAAGAAATGCCCCAGA 480

QY 543 TATTCAGAAAGCTGAACCAAGATGGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGG 602
Db 481 TATTCAGAAAGCTGAACCAAGATGGAAGATCATATTTACTGAGATACCTTCAATCCAGTTGG 540

QY 603 GGTCTATGAAGTGGCAAACTCGGTGAAGATCCC 636
Db 541 GGTCTATGAAGTGGCAAACTCGGTGAAGATCCC 574

RESULT 7

AW472274
LOCUS
DEFINITION
si23b06.y1 Gm-c1029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1029-2028 5' similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE 1; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

REFERENCE
AUTHORS

Glycine.
1 (bases 1 to 567)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

TITLE
JOURNAL
COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63109, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1331 Std Error: 0.00
High quality sequence stop: 426.

FEATURES
source

1..567
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1029-2028"
/tissue_type="very young cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-c1029"
/notes="Vector: pSPORT1; Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mg fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

ORIGIN

Query Match 42.9%; Score 562.8; DB 10; Length 567;
Best Local Similarity 99.8%; Pred. No. 1.4e-86;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 67 GGTGGTCCCGGTTTCATTGGCACCCACACCGTGTTCAGCTTCTCAAAGCTGGCTTCAGC 126
Db 1 GGTGGTCCCGGTTTCATTGGCACCCACACCGTGTTCAGCTTCTCAAAGCTGGCTTCAGC 60

QY 127 GTTTCATATATCGACAATTTGGATCTCCGTCATGAGACGATGACCGCGTCCGCCAA 186
Db 61 GTTTCATATATCGACAATTTGGATCTCCGTCATGAGACGATGACCGCGTCCGCCAA 120

QY 187 GTGGTGGCCCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGGCGGATCTCCGGAATAGG 246
Db 121 GTGGTGGCCCTCTGCTTTCTCAGAACCTCCCAATTCACCCAGGCGGATCTCCGGAATAGG 180

QY 247 GATGACTTGGAGAACTCTTCTCCAAAACAATTTGATGCGGTGATCCACTTTGTGGC 306
Db 181 GATGACTTGGAGAACTCTTCTCCAAAACAATTTGATGCGGTGATCCACTTTGTGGC 240

QY 307 TTGAAAGCGGTGCTGCAAAACGTTGCGAACCCCGCTCGCTATTTTGAATTTAATTTGGTT 366
Db 241 TTGAAAGCGGTGCTGCAAAACGTTGCGAACCCCGCTCGCTATTTTGAATTTAATTTGGTT 300

QY 367 GGCACCATCAACCTCTACGAGTTTATGGCAAGTATATATGCAAAAGATGGTTTCTCA 426

FEATURES	Source
Location/Qualifiers	
1..564	
/organism="Glycine max"	
/mol_type="mRNA"	
/db_xref="taxon:3847"	
/clones="GENOME SYSTEMS CLONE ID: Gm-r1030-3616"	
/lab_host="DH10B"	
/clone_lib="Gm-r1030"	
/note=vector: pSPORT1; Site 1: SalI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI	

JOURNAL Unpublished (2001)

JOURNAL Unpublished (2001)

TOITPNAT

COMMENT

Contact: Michael G. Hahn
Complex Carbohydrate Research Center
University of Georgia
220 Riverbend Road, Athens, GA 30602-4712, USA
Tel: 706-542-4457
Fax: 706-542-4412
Email: hahn@ccrc.uga.edu
TIGR sequence name: MTMCW07TK
More information is available at: www.medicago.org
Seq primer: Slnod (CTA gaa cta gta gat cc).

FEATURES

source
1..798
Location/Qualifiers

/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="HOGA-19B13"
/tissue_type="3 day old seedling roots"
/dev_stage="24 hours after treatment in the dark at 26 C
with 0.5 mg/ml oligogalacturonides (DP 6-20) in the
presence of 100 ug/ml Gentamicin"
/lab_host="XLOLR"
/clone_lib="HOGA"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in SOLR cells."

ORIGIN

Query Match 42.7%; Score 560.2; DB 14; Length 798;
Best Local Similarity 83.3%; Pred. No. 3.4e-86;
Matches 637; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 13 TCTCTCCTCTATTGCGACGATGGTCTCTCCCAACACATCTCGTCAACCGGTGT 72
DB 34 TCTCTGTTCCAAATCCAAAATGGTTTCACTTCCCAAGATCTCGTCAACCGGTGT 93

QY 73 GCGCGTTTCAATGGCCACCAACCGTGTTCAGCTTCTCAAAGCTGGCTTCAGCGTTCA 132
DB 94 TCTGTTTTCATCGGCACTCACACGGTGTTCAGCTTCTCAAAGTGGTTCGTCTCC 153

QY 133 ATATCGCAATTCGATTAATCTCGTCATGAGCAGTGGACCGTCCGCCAAGTGT 192
DB 154 ATTATTGATAATTCGATAATCTGTATAGACGCTGTGGACCGTTCGTGAATGT 213

QY 193 GGCCTCTGCTTCTCAGAACCTCCCAATTCACCCAGGCGATCTCCGGAATAGGATGAC 252
DB 214 GGTCTCAACTTCTCAGATCTCGATTTCACATTTGGGAGATCTCAGATTAAGATGAT 273

QY 253 TTGAGAACTCTTCTCCAAAACAACATTTGATGCGGTGATCCACTTTCGCTGTGAA 312
DB 274 TTGAGAACTCTTCTTAAACTTAAATTCGACCGGTGATTCATTTGCTGGATTAAA 333

QY 313 GCGCTTCTCAAGCGTTGGCAAGCCCGTGTCTATTTTGAATTTAATTTGGTGGCAC 372
DB 334 GCGCTTGGTGAAGTGTGGATCCCGCTGTCTATTTGATAATCTTGTGGACT 393

QY 373 ATCAACTCTACAGTTTATGGCAAGATTAATGTAATGCAAAAGATGGTTTTCTCATCT 432
DB 394 ATCAATCTTATGAAGTTATGGCAATATAATTTGAAAAGATGGTTTTCTCATCT 453

QY 433 GCACCGTTTATGGCAACTGAAAGATACCGTGTGAGAGATTTCAAGTTACAGCT 492
DB 454 GCTACTGTTTATGTTCACTGACGATACCTGTGTGGAGATTTCAAGTTACAGCT 513

QY 493 ATGAATCCCTATGAGCGCAACAGCTTTTCTCGAAGAAATTCGCCGAGATTTAGAAA 552
DB 514 ATGAATCCCTATGCGCGCACTAAGCTTTTCTTGAAGAAATTCGAGAGATTTAGATA 573

QY 553 GCTGAACCAAGATGAGAGATCATATTACTAGATACCTTCATCCAGTGGGGTCA 612

DB 574 GCTGAACCAAGATGAGAGATCATATTACTGAGATCACTTCAATCCAGTTGGAGCCCATGAA 633
QY 613 AGTGGCAAACTCGGTGAAGATCCCAAGGCGATCCCAATAACCTCATGCTTACATCAG 672
DB 634 AGTGGTAAACTCGGTGAAGATCCCAAGGCGATCCCAATAATCTCATGCTTATACAT 693
QY 673 CAAGTAGCTGTGGAAGATTTGACTGAACCTCAATGTATACGGTCAATGATTCACCAAGGAG 732
DB 694 CAAGTAGCTGTGGAAGATTTGACTGAACCTCGGCTCAATGTATATGGTCAATGATTCATCTACAGG 753

QY 733 GATGGCTCTCGGATCCGGACTATATCCATGTGATGACATTGGCA 777
DB 754 GATGGCTCTCGGTTCCGGACTATATCCATGTGATGACTTAGCA 798

RESULT 10

AWS07982

LOCUS

DEFINITION

SOURCE

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..580

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clones="GENOME SYSTEMS CLONE ID: Gm-r1030-2084"

/lab_host="DH10B"

/clone_lib="Gm-r1030"

/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This

cDNA library was constructed from mRNA isolated from

immature cotyledons of greenhouse grown plants

(individual seed fresh weight of 100-300mg). The library

was prepared using the Life Technologies pSuperScript cDNA

library construction kit. Complementary DNA was

synthesized from mRNA using a poly(dT) sequence with a

NotI restriction site. SalI linkers adapters were ligated

to the blunt-ended cDNA fragments followed by NotI

digestion. The cDNA fragments were directionally cloned

into the NotI-SalI restriction site of the pSPORT1

vector. The ligated cDNA fragments were transformed into

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was

E. coli ElectroMax DH10B host cells. This library was


```

QY 488 AAGCTATGATCCCTATGACGACGACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTC 547
Db 482 AAGCTATGATCCCTATGACGACGACCAAGCTTTCTCGAAGAAATTCGCCGAGATATTC 541
QY 548 AGAAGCTGAACGAGATGGA 568
Db 542 AGAAGCTGAACGAGATGGA 562

RESULT 12
LOCUS BE474654
DEFINITION sp66910.y1 Gm-cl044 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl044-955 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ;, mRNA sequence.
ACCESSION BE474654
VERSION BE474654.1 GI:9565136
KEYWORDS EST
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 560)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanma,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ocu@resgen.com
Insert Length: 1398 Std Error: 0.00
High quality sequence stop: 347.
FEATURES
source Location/Qualifiers
1..560
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl044-955"
/tissue_type="Hypocotyl, 9-10 day old etiolated seedlings"
/lab_host="DH108"
/clone_lib="Gm-cl044"
/notes="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from etiolated hypocotyl tissue of 9-10 day old seedlings
of the cultivar Williams. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(GT) primer with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
followed by digestion with EcoRI and XhoI. The cDNA
fragments were directionally cloned into the EcoRI-XhoI
restriction site of the pBluescript vector. The ligated
cDNA fragments were transformed into DH10B host cells
(Gibco BRL). This library was constructed by Dr. Randy
Shoemaker."

ORIGIN
Query Match 42.0%; Score 550.8; DB 10; Length 560;
Best Local Similarity 99.6%; Pred. No. 1.6e-84;
Matches 552; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 CCACCTCTCTCCCTCTCTATTCGAGCATGGTGTCTTCTCTCCCAACACATTCCTGGTCACCG 67
Db 7 CCACCTCTCTCCCTCTCTATTCGAGCATGGTGTCTTCTCTCCCAACACATTCCTGGTCACCG 66
QY 68 GTGGTGGCCGCTTTCATTGGCACCACCAACCGTGGTTCAGCTTCTCAAGAGTGGCTTCAGCG 127
Db 67 GTGGTGGCCGCTTTCATTGGCACCACCAACCGTGGTTCAGCTTCTCAAGAGTGGCTTCAGCG 126
QY 128 TTTCATAATTCGACAAATTCGATAAATCCCGTCAATGAAGCAGTGGACCCGCTCCGCCAAG 187
Db 127 TTTCATAATTCGACAAATTCGATAAATCCCGTCAATGAAGCAGTGGACCCGCTCCGCCAAG 186
QY 188 TGGTGGCCCTCTGCTTCTCAGAACTCCAATTCACCCAGGCGCATCTCCGGAATAGGG 247
Db 187 TGGTGGCCCTCTGCTTCTCAGAACTCCAATTCACCCAGGCGCATCTCCGGAATAGGG 246
QY 248 ATGACTTGGAGAAACTTCTTCTCCAAAACAACTTGTATGATCCGCTGATCCACTTTCGGCT 307
Db 247 ATGACTTGGAGAAACTTCTTCTCCAAAACAACTTGTATGATCCGCTGATCCACTTTCGGCT 306
QY 308 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGCTGCTATTTTGAATTTAAATTTGGTTG 367
Db 307 TGAAGCGGTGCTGAAAGCGTTGCGAAGCCCGCTGCTATTTTGAATTTAAATTTGGTTG 366
QY 368 GCACCATCAACCTCTACGAGTTTATGCGAAAGTATAATTGCAAAAAGATGGTTTTCTCAT 427
Db 367 GCACCATCAACCTCTACGAGTTTATGCGAAAGTATAATTGCAAAAAGATGGTTTTCTCAT 426
QY 428 CATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAGGATTTCAAGTTAC 487
Db 427 CATCTGCAACCGTTTATGCGCAACCTGAAAGATACCGTGTGAGGAGGATTTCAAGTTAC 486
QY 488 AAGCTATGAATCCCTATGAGCGGACCAAGCTTTTCTCGAAGAAATTCGCCGAGATATTC 547
Db 487 AAGCTATGAATCCCTATGAGCGGACCAAGCTTTTCTCGAAGAAATTCGCCGAGATATTC 546
QY 548 AGAAGCTGAACCA 561
Db 547 AGAAGCTGAACCA 560

RESULT 13
LOCUS BM528276
DEFINITION sal56d10.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl061-2325 5', similar to TR:Q43070 Q43070
UDP-GALACTOSE-4-EPIMERASE ;, mRNA sequence.
ACCESSION BM528276
VERSION BM528276.1 GI:18734601
KEYWORDS EST
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 547)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
Khanma,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130

```

South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cu@resgen.com web site:

www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES

Location/Qualifiers

1. .547
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-3235"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl061"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 41.2%; Score 540.6; DB 12; Length 547;
Best Local Similarity 99.3%; Pred. No. 9.1e-83;
Matches 543; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 627 TGAAGATCCCAAGGCGATCCCAATACCTCATGCTTACCTTACATTCACGACAGTAGCTGTGG 686
Db 1 TGAAGATCCCAAGGCGATCCCAATACCTCATGCTTACCTTACATTCACGACAGTAGCTGTGG 60
Qy 687 AAGATTGACGAACTCAATGATACGGTCAATGATATCCAAAGAGGATGGCTCGCAT 746
Db 61 AAGATTGACGAACTCAATGATACGGTCAATGATATCCAAAGAGGATGGCTCGCAT 120
Qy 747 CCGGACTATATCCATGTGATGACATTCGGCAGATGCCATATTGCTGCCCTCGCAAGCT 806
Db 121 CCGGACTATATCCATGTGATGACATTCGGCAGATGCCATATTGCTGCCCTCGCAAGCT 180
Qy 807 CTTCCAGACGAGACATAGTTGTACTGCTTACACCTGGGACCTGGTGGACATC 866
Db 181 CTTCCAGACGAGACATAGTTGTACTGCTTACACCTGGGACCTGGTGGACATC 240
Qy 867 TGTGCTTGAATGGTTACAGCAITTTGAAAGGCTTCTGGCAAGAAATTCAGTAAAT 926
Db 241 TGTGCTTGAATGGTTGACAGATTTGAAAGGCTTCTGGCAAGAAATTCAGTAAAT 300
Qy 927 ATGTCGAAGAACCGGAGATCGGACTGAGTTTATGCACTACAGAGAGCTGAGAA 986
Db 301 ATGTCGAAGAACCGGAGATCGGACTGAGTTTATGCACTACAGAGAGCTGAGAA 360
Qy 987 AGAATTTGTTGGAAGCAAACTATGTTGTGGAGGAGATGTGACGAGCAACCAATGGAATTG 1046
Db 361 ATAACTTGTGTTGAGGCAAACTATGTTGTGGAGGAGATGTGACGAGCAACCAATGGAATTG 420
Qy 1047 GGCAGAGAAACAATCCCTGGGGTTACCGGGGAGAGCTTGAATGCTTACAGAAATATCT 1106
Db 421 GGCAGAGAAACAATCCCTGGGGTTACCGGGGAGAGCTTGAATGCTTACAGAAATATCT 480
Qy 1107 GCTCATCTACGATGCTTTTACATAAATAGGATCTCTTATATAGATACCTTTTATGTT 1166
Db 481 GCTCATCTACGATGCTTTTACATAAATAGGATCTCTTATATAGATACCTTTTATGTT 540
Qy 1167 TGATGAT 1173
Db 541 GGATGAT 547

RESULT 14

BF635019

LOCUS

DEFINITION

BF635019

ACCESSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 694)

AUTHORS

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula drought library

Unpublished (2000)

JOURNAL

COMMENT

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7389

Email: gdmay@noble.org

Insert Length: 694 Std Error: 0.00

Plate: 078 row: G column: 09

Seq primer: TCACACAGGAAACACGATGAC.

Location/Qualifiers

1. 694

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NF07809DT"

/tissue_type="Plantlets"

/dev_stage="Pooled timepoints"

/clone_lib="Drought"

/note="Vector: Lambda Zap; Contains a mixture of entire

plantlets harvested in a series of days-post-watering

timepoints."

ORIGIN

Query Match 40.8%; Score 535.2; DB 10; Length 694;
Best Local Similarity 85.6%; Pred. No. 6.9e-82;
Matches 594; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 143 ATTTCGATAATCTCGTCATGGAAGAGTGGACCCGCTCGCCCAAGTGGTTGGCCCTCTGC 202
Db 1 ATTTTCGATAATCTCGTCATGGAAGAGTGGACCCGCTCGCCCAAGTGGTTGGCCCTCTGC 60
Qy 203 TTTCTCAGAACTCCAAATTCACCCAGGCGGATCTCCGGAATAGGATGACCTGGAGAAAC 262
Db 61 TTTCTCAGAACTCCAAATTCACCTTTGGAGATCTCAGAAATTAAGATGATTTGGAGAAAC 120
Qy 263 TCTTCTCCAAACCAACATTTGATCCCGTGATCCACTTTGCTGGCTTGAAGCGGTGCTG 322
Db 121 TCTTTTCTAAACTAAATTCGACCGGTGATTCATTTTCTGCGATTAAAGCCGTTGGTG 180
Qy 323 AAAGGTTTGGCAAGCCCGCTGCTATTTTGATTTTAAATTTGGTGGCACCACCACTCT 382
Db 181 AAAGTGTTCGAAATCCCGCTGCTATTTTGATAATAATCTTTTGGGACTATCAATCTTT 240
Qy 383 ACGAGTTTATGGCAAGATATAATTCGCAAAAAGATGGTTTTCTCATCATCTCAACCGTTT 442
Db 241 ATGAAGTTATGGCCAAATATAATTTGTAATAAGATGGTTTTTTCATCATCTCTACTGTT 300
Qy 443 ATGCCAACTGAAAGATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCT 502
Db 301 ATGSTCAACCTGACACGATACCCCTGTGTGGAGGATTTCAAGTTACAGCCATGATCCCT 360
Qy 503 ATGGACGGACCAAGCTTTTCTCTGGAAGAAATTTGCCGAGATATTTCAGAAAGCTGACCAAG 562

Db 361 ATGGCGGAGTAAGCTTTTCCCTTGAAGAAATTCGAGAGATATTCAGATAGCTGACCG 420
 QY 563 AATGGAAGATCATATTACTGAGATATTCATATCCAGTTGGGCTCATGAAAGTGCACAAAC 622
 Db 421 AATGGAAGATCATATTACTGAGATATTCATATCCAGTTGGGCTCATGAAAGTGCACAAAC 480
 QY 623 TGGGTGAAGATCCCAAGGATCCCAATAAATCACTATGCTTACATTCAGCAAGTACGCTG 682
 Db 481 TGGGTGAAGATCCCAAGGATCCCAATAAATCACTATGCTTACATTCAGCAAGTACGCTG 540
 QY 683 TTGGAAGATGATGCTGAATCAATGTATAGCTCATGATTCATCAACGAGGATGCTCTG 742
 Db 541 TTGGAAGATGATGCTGAATCAATGTATAGCTCATGATTCATCAACGAGGATGCTCTG 600
 QY 743 CGATCCGGAATATATCCATGTGATGAGTTCGAGATGGCCATATTCCTGCTCGCAA 802
 Db 601 CGGNTCCGGAATATATCCATGTGATGAGTTCGAGATGGCCATATTCCTGCTCGCAA 660
 QY 803 AGCTCTTCAACCGGAGACATAGGTTGTACTGC 836
 Db 661 AGCTTTTCGACGAGGAACATANGTTGTACTGC 694

RESULT 15

BM891597 552 bp mRNA linear EST 11-MAR-2002
 LOCUS sam41d04.y1 Gm-c1068 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-c1068-7447 5' similar to TR:Q43070 Q43070
 DEFINITION UDP-GALACTOSE-4-EPIMERASE ; mRNA sequence.

ACCESSION

BM891597

VERSION

EST.

KEYWORDS

Glycine max (soybean)

SOURCE

Glycine max

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCaun, R., Waterston, R. and Willson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 426.

Location/Qualifiers

1. .552

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1068-7447"

/tissue_type="Leaf, drought stressed, 1 month old plants,

greenhouse grown"

/lab host="DH10B"

/clone lib="Gm-c1068"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from drought stressed leaf tissue of the cultivar Williams
 82. The month old greenhouse grown plants were deprived of
 water for 3 days prior to harvesting the stressed leaf
 tissue. Complementary DNA was synthesized from mRNA using
 a primer consisting of a poly(dT) sequence with a XhoI
 restriction site. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the pBluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 40.5%; Score 532; DB 12; Length 552;
 Best Local Similarity 98.7%; Pred. No. 2.7e-81;
 Matches 546; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
 QY 39 GTCTTCCTCCCAACACATTTCTGTCACCGTGTGCGGTTTCATTTGGGACCCACACCGT 98
 Db 1 GTCTTCCTCCCAACACATTTCTGTC-CCGATGTGCGGTTTCATTTGGGACCCACACCGT 59
 QY 99 CGTTTCAGCTTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAATTTTCGATAACTCCGT 158
 Db 60 CGTTTCAGCTTTCTCAAGCTGGCTTCAGCGTTTCAATAATCGACAATTTTCGATAACTCCGT 119
 QY 159 CATGGAAGCAGTGACCGCTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCA 218
 Db 120 CATGGAAGCAGTGACCGCTCCGCAAGTGGTGGCCCTCTGCTTCTCAGAACCTCCA 179
 QY 219 ATTACACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCCAAACAAC 278
 Db 180 ATTACACCCAGGCGATCTCCGGAATAGGATGACTTGGAGAAACTCTTCTCCAAACAAC 239
 QY 279 ATTTGATGCCGTGATCCATTTGCTGGCTTGAAGCGGTTGCTGAAGCGTTGGGAAGCC 338
 Db 240 ATTTGATGCCGTGATCCATTTGCTGGCTTGAAGCGGTTGCTGAAGCGTTGGGAAGCC 299
 QY 339 CCGTCGCTATTTGATTTTAAATTTGGTTGGCACCCTCACTCTACGAGTTTATGGCAA 398
 Db 300 CCGTCGCTATTTGATTTTAAATTTGGTTGGCACCCTCACTCTACGAGTTTATGGCAA 359
 QY 399 GTATAATTGCAAAAAGATGTTTTCTCATCATCTGCAACCGTTTATGGCAACTGAAAA 458
 Db 360 GTATAATTGCAAAAAGATGTTTTCTCATCATCTGCAACCGTTTATGGCAACTGAAAA 419
 QY 459 GATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTATGACGCGACCAAGCT 518
 Db 420 GATACCGTGTGAGGAGGATTTCAAGTTACAGCTATGATCCCTATGACGCGACCAAGCT 479
 QY 519 TTTCTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAAACCGAATGGAAGATCATATT 578
 Db 480 TTTCTCTGGAAGAAATTCGCCGAGATATTCAGAAAGCTGAAACCGAATGGAAGATCATATT 539
 QY 579 ACTGAGATATCTC 591
 Db 540 ACTGAGAACTTC 552

Search completed: March 22, 2004, 20:17:57

JOB time : 3776 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 20:18:05 ; Search time 74 Seconds
(without alignments)
1336.373 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSQHLVITGGAGFIGTH.....MCRDQWNAKNPWGACKP 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1845	100.0	350	3 AAB07819	AAB07819 Amino aci
2	1530	82.9	351	3 AAG09987	AAG09987 Arabidops
3	1486	80.5	351	5 AAG79862	AAG79862 Potato DG
4	1384	75.0	352	3 AAB07820	AAB07820 Amino aci
5	1358	73.6	339	3 AAG15804	AAG15804 Arabidops
6	1231	66.7	375	3 AAG15222	AAG15222 Arabidops
7	1231	66.7	377	3 AAG42840	AAG42840 Arabidops
8	1229	66.6	348	3 AAG15224	AAG15224 Arabidops
9	1229	66.6	349	3 AAG15223	AAG15223 Arabidops
10	1229	66.6	350	3 AAG42842	AAG42842 Arabidops
11	1229	66.6	351	3 AAG42841	AAG42841 Arabidops
12	1222	66.2	348	3 AAG53259	AAG53259 Arabidops
13	1222	66.2	372	3 AAG53258	AAG53258 Arabidops
14	1220	66.1	348	3 AAG21995	AAG21995 Arabidops
15	1220	66.1	372	3 AAG21994	AAG21994 Arabidops
16	1208	65.5	350	3 AAG51824	AAG51824 Arabidops
17	1201	65.1	350	3 AAG42204	AAG42204 Arabidops
18	1192	64.6	349	3 AAB07823	AAB07823 Amino aci
19	1180	64.0	354	3 AAB07822	AAB07822 Amino aci
20	1094	59.3	231	3 AAG09988	AAG09988 Arabidops
21	1054	57.2	254	3 AAB07821	AAB07821 Amino aci
22	1054	57.1	224	3 AAG09989	AAG09989 Arabidops
23	1038	56.3	339	4 AAG64602	AAG64602 Uridine d
24	1038	56.3	339	5 AAB881469	AAB881469 B. subtil
25	1012.5	54.9	266	3 AAG53260	AAG53260 Arabidops

26	1010.5	54.8	266	3 AAG21996	AAG21996 Arabidops
27	986.5	53.5	348	3 AAY54118	AAY54118 Amino aci
28	986.5	53.5	348	5 ABG96567	ABG96567 Human sho
29	977.5	53.0	348	2 AAW01619	AAW01619 Human uri
30	977.5	53.0	348	4 AAB51152	AAB51152 Human UDP
31	977.5	53.0	348	7 ADA66613	ADA66613 Human uri
32	966	52.4	339	6 ABU02070	ABU02070 S. pneumo
33	952	51.6	341	6 ABM68787	ABM68787 Phototrab
34	941	51.0	336	6 ABU02313	ABU02313 S. pneumo
35	927	50.2	340	5 ABP66248	ABP66248 Bifidobac
36	918.5	49.8	350	4 ABB57926	ABB57926 Drosophil
37	917.5	49.7	341	6 ADB09548	ADB09548 Alloiococ
38	904.5	49.0	338	5 ABB81470	ABB81470 N. gonorr
39	904.5	49.0	340	6 ABP80860	ABP80860 N. gonorr
40	904.5	49.0	340	6 ABP80439	ABP80439 N. gonorr
41	898	48.7	219	3 AAG15805	AAG15805 Arabidops
42	893.5	48.4	338	3 AAY54117	AAY54117 Amino aci
43	891.5	48.3	338	3 AAY81359	AAY81359 Escherich
44	887.5	48.1	235	3 AAG51825	AAG51825 Arabidops
45	878.5	47.6	699	5 ABG93041	ABG93041 S. cerevi

ALIGNMENTS

RESULT 1

AAB07819
ID AAB07819 standard; protein; 350 AA.

AC AAB07819;

XX 14-NOV-2000 (first entry)

XX Amino acid sequence of UDP-galactose 4-epimerase.

XX UDP-galactose 4-epimerase; UDP-glucose; carbon flux;
KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level.

OS Glycine max.

XX Key Location/Qualifiers

FT Misc-difference 3..129
FT /note= "the nucleotides encoding these residues are not given"

XX WO200047755-A2.

XX 17-AUG-2000.

XX 09-FEB-2000; 2000WO-US003453.

XX 10-FEB-1999; 99US-0119588P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PI Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;

PI Rafalski JA, Thorpe CJ;

XX WPI; 2000-549152/50.

XX N-ESDB; AAA59439.

XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant cells to e.g. decrease cell wall constituents and increase starch content.

XX Claim 10; Page 51-52; 61pp; English.

XX The present sequence represents a UDP-galactose 4-epimerase. The enzyme is responsible for the interconversion of UDP-glucose and UDP-galactose. As the activity of the enzyme appears to be particularly limiting to carbon flux into the raffinoseaccharide pathway, reduction of its activity should decrease the levels of raffinose and stachyose in seeds. The

PR 01-FEB-2001; 2001US-0265311P.
XX (DAIZ/) DAI Z.
XX Dai Z, Shi L, Hooker BS;
XX WPI; 2002-598747/75.
XX N-PSDB; ABA00335.
XX New isolated polynucleotide molecule comprising a gene coding for the
PT enzyme 5'-diphospho galactose 4-epimerase, useful for regulating
PT carbohydrate metabolism or nutritional profile in transgenic plants.
XX
XX Claim 1; Fig 1; 38pp; English.
XX This sequence is encoded by the potato psen-1 gene and represents the
CC enzyme 5'-diphospho-galactose 4-epimerase (DGE). This sequence is
CC homologous to the sequence of Arabidopsis thaliana uridine diphosphate
CC glucose (UDPG) epimerase gene (86% homology at peptide level, 71%
CC homology at the nucleotide level). The DGE polynucleotide molecules are
CC useful in regulating carbohydrate metabolism in transgenic plants. The
CC polynucleotide can be used in Controlled Environment Agriculture that
CC employs an integrated system for commercial production of transgenic
CC plants in a controlled environment. Controlling the level of DGE in
CC transgenic plants is useful for regulating the nutritional profile of the
CC plant
XX Sequence 351 AA;
SQ
Query Match 80.5%; Score 1486; DB 5; Length 351;
Best Local Similarity 78.2%; Pred. No. 3.3e-133;
Matches 273; Conservative 33; Mismatches 43; Indels 0; Gaps 0;
QY 2 VSSSOHLVTGGAGTGTHTVQLLKAGFSVSIIDNFNSVMEAVDRVQVGPQLLSQNL 61
Db 3 VQCCENILVTGGAGTGTHTVQLNEGPKVTIIDNFNSVEAVDRVRELVPQLSQNL 62
QY 62 QFTQGLRRDDLEKLFSTTFDAVTHFAGLKAVAESVAKPRYPDFNLVGTINLYEFMA 121
Db 63 EPHLGDIRKDDLEKLFSEKFAAVVHFAGLKAVGESVQVPELYFENNIGSITLYSMA 122
QY 122 KYNCKMVPSSSATVYGQPEKIPCEEDFKLQAMNPYGRTKLFLLEETARDIOKAEPEWII 181
Db 123 KYNCKLVPSSSATVYGQPEKIPCEEDFKLQAMNPYGRTKLFLLEETARDIOKAEPEWII 182
QY 182 LLRYFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNHYGHDYPTDGSAIRDY 241
Db 183 LLRYFNPVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLPELVNHYGNDYPTDGTAIRDY 242
QY 242 IHVMDLADGHIALRLFTTENIGCTAYNLGRTGTSVLEMTAFKASGKKIPVKLCPR 301
Db 243 IHVLDLADGHVVALQRLLRQNLHGLGCVAYNLGTGKSKVLEWVAAPERASGKKIPLKMCPR 302
QY 302 RPDGATEVYASTERAESKELGWKANYGVEMCRDQWNAKNPVGAGKP 350
Db 303 RPDGATVYASTEAEKELGWAKYGINEMCRDQWNAKNPVGQSKP 351
RESULT 4
AAB07820
ID AAB07820 standard; protein; 352 AA.
XX
XX AAB07820;
XX
DT 14-NOV-2000 (first entry)
XX Amino acid sequence of UDP-galactose 4-epimerase.
XX
XX UDP-galactose 4-epimerase; UDP-glucose; UDP-galactose; carbon flux;
KW raffinoseaccharide pathway; raffinose; stachyose; carbohydrate metabolism;
KW starch level.
XX
XX Triticum aestivum.
OS

XX W0200047755-A2.
XX 17-AUG-2000.
XX 09-FEB-2000; 2000WO-US003453.
XX 10-FEB-1999; 99US-0119588P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Carlson TJ, Fader GM, Famodu OO, Kinney AJ, Pearlstein RW;
XX Rafalski JA, Thorpe CJ;
XX WPI; 2000-549152/50.
XX N-PSDB; AAA59440.
XX Novel DNA encoding a UDP-galactose-4-epimerase used to transform plant
PT cells to e.g. decrease cell wall constituents and increase starch
PT content.
XX Claim 10; Page 52-53; 61pp; English.
XX The present sequence represents a UDP-galactose 4-epimerase. The enzyme
CC is responsible for the interconversion of UDP-glucose and UDP-galactose.
CC As the activity of the enzyme appears to be particularly limiting to
CC carbon flux into the raffinoseaccharide pathway, reduction of its activity
CC should decrease the levels of raffinose and stachyose in seeds. The
CC polynucleotide is useful to transform plant cells to enhance the degree
CC of expression of UDP-galactose-4-epimerase within these cells. Changes in
CC the expression of UDP-galactose-4-epimerase within a cell enables
CC modification of plant carbohydrate metabolism. This allows growth of
CC grains with reduced cell wall constituents (fiber) and increased levels
CC of starch
XX Sequence 352 AA;
SQ
Query Match 75.0%; Score 1384; DB 3; Length 352;
Best Local Similarity 73.6%; Pred. No. 1.9e-123;
Matches 251; Conservative 42; Mismatches 48; Indels 0; Gaps 0;
QY 8 ILVTGGAGTGTHTVQLLKAGFSVSIIDNFNSVMEAVDRVQVGPQLLSQNLQFTQGD 67
Db 5 VLVTGGAGTGTHTVQLLEKGYAVTAVDNFNSVPEALDRVRHIVGPALSARLQIFGD 64
QY 68 LRNRDDLEKLFSTTFDAVTHFAGLKAVAESVAKPRYPDFNLVGTINLYEFMAKNCKK 127
Db 65 LTIKDDLEKVFAAKQYDAVTHFAGLKAVAESVAHPENYRNNIVGTNLYDVMKGGCNK 124
QY 128 MVSSSATVYGQPEKIPCEEDFKLQAMNPYGRTKLFLLEETARDIOKAEPEWIIILRYFN 187
Db 125 LVFSSSATVYGQPEKIPCEEDFKLQAMNPYGRTKLFLLEETARDIOKAEPEWIIILRYFN 184
QY 188 PVGAHESGKLGEDPKGIPNNLMFYIQQAVAGRLTELVNHYGHDYPTDGSAIRDYIHVMDL 247
Db 185 FIGAHESGDIQEDPKGIPNNLLFYIQQAVARRPELVNHYGHDYPTDGTAVRDYIHVMDL 244
QY 248 ADGHIALRLFTTENIGCTAYNLGRTGTSVLEMTAFKASGKKIPVKLCRRRPGDAT 307
Db 245 ADGHIAALEKLFATPDIGCVAYNLGTGRTGTSVLEWVAFAKGYKKIPVKLCRRRPGDSE 304
QY 308 EVYASTERAEKELGWKANYGVEMCRDQWNAKNPVGAG 348
Db 305 QVYASTAKAEELGWAKYGIEMCRDQWNAKNPVGAG 345
RESULT 5
AAG15804
ID AAG15804 standard; protein; 339 AA.
XX
XX AAG15804;
XX
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 16201.
DE XX
XX XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-00301439.
XX XX
XX 25-FEB-1999; 99US-0121825P.
XX PR 22-JUN-1999; 99US-0139899P.
XX PR 23-JUN-1999; 99US-0140353P.
XX PR 23-JUN-1999; 99US-0140354P.
XX PR 24-JUN-1999; 99US-0140659P.
XX PR 28-JUN-1999; 99US-0140823P.
XX PR 29-JUN-1999; 99US-0140991P.
XX PR 30-JUN-1999; 99US-0141287P.
XX PR 01-JUL-1999; 99US-0141842P.
XX PR 01-JUL-1999; 99US-0142154P.
XX PR 02-JUL-1999; 99US-0142055P.
XX PR 06-JUL-1999; 99US-0142390P.
XX PR 08-JUL-1999; 99US-0142803P.
XX PR 09-JUL-1999; 99US-0142920P.
XX PR 12-JUL-1999; 99US-0142977P.
XX PR 13-JUL-1999; 99US-0143542P.
XX PR 14-JUL-1999; 99US-0143624P.
XX PR 15-JUL-1999; 99US-0144005P.
XX PR 16-JUL-1999; 99US-0144085P.
XX PR 16-JUL-1999; 99US-0144086P.
XX PR 19-JUL-1999; 99US-0144325P.
XX PR 19-JUL-1999; 99US-0144331P.
XX PR 19-JUL-1999; 99US-0144332P.
XX PR 19-JUL-1999; 99US-0144333P.
XX PR 19-JUL-1999; 99US-0144334P.
XX PR 19-JUL-1999; 99US-0144335P.
XX PR 20-JUL-1999; 99US-0144352P.
XX PR 20-JUL-1999; 99US-0144632P.
XX PR 20-JUL-1999; 99US-0144884P.
XX PR 21-JUL-1999; 99US-0144814P.
XX PR 21-JUL-1999; 99US-0145086P.
XX PR 21-JUL-1999; 99US-0145088P.
XX PR 22-JUL-1999; 99US-0145085P.
XX PR 22-JUL-1999; 99US-0145087P.
XX PR 22-JUL-1999; 99US-0145089P.
XX PR 22-JUL-1999; 99US-0145152P.
XX PR 23-JUL-1999; 99US-0145145P.
XX PR 23-JUL-1999; 99US-0145218P.
XX PR 23-JUL-1999; 99US-0145224P.
XX PR 26-JUL-1999; 99US-0145276P.
XX PR 27-JUL-1999; 99US-0145913P.
XX PR 27-JUL-1999; 99US-0145918P.
XX PR 27-JUL-1999; 99US-0145919P.
XX PR 28-JUL-1999; 99US-0145951P.
XX PR 02-AUG-1999; 99US-0146386P.
XX PR 02-AUG-1999; 99US-0146388P.
XX PR 02-AUG-1999; 99US-0146389P.
XX PR 03-AUG-1999; 99US-0147038P.
XX PR 04-AUG-1999; 99US-0147204P.
XX PR 04-AUG-1999; 99US-0147302P.
XX PR 05-AUG-1999; 99US-0147192P.
XX PR 05-AUG-1999; 99US-0147260P.
XX PR 06-AUG-1999; 99US-0147303P.
XX PR 06-AUG-1999; 99US-0147416P.
XX PR 09-AUG-1999; 99US-0147492P.
XX PR 09-AUG-1999; 99US-0147935P.
XX PR 10-AUG-1999; 99US-0148171P.
XX PR 11-AUG-1999; 99US-0148319P.
XX PR 12-AUG-1999; 99US-0148341P.
XX PR 13-AUG-1999; 99US-0148565P.
XX PR 13-AUG-1999; 99US-0148684P.
XX PR 16-AUG-1999; 99US-0149368P.
XX PR 17-AUG-1999; 99US-0149175P.
XX PR 18-AUG-1999; 99US-0149426P.
XX PR 20-AUG-1999; 99US-0149722P.
XX PR 20-AUG-1999; 99US-0149723P.
XX PR 20-AUG-1999; 99US-0149929P.
XX PR 23-AUG-1999; 99US-0149902P.
XX PR 23-AUG-1999; 99US-0149930P.
XX PR 25-AUG-1999; 99US-0150566P.
XX PR 26-AUG-1999; 99US-0150884P.
XX PR 27-AUG-1999; 99US-0151065P.
XX PR 27-AUG-1999; 99US-0151066P.

QY	303 PGD 305	304 SGD 306
Db		
RESULT 6		
AAG15222		
ID	AAG15222 standard; protein; 375 AA.	
XX	AC	AAG15222;
XX	17-OCT-2000	(first entry)
XX	Arabidopsis thaliana protein fragment	SEQ ID NO: 15390.
DE	Arabidopsis thaliana	
XX	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	OS	Arabidopsis thaliana.
XX	PN	EP1033405-A2.
XX	PD	06-SEP-2000.
XX	25-FEB-2000;	2000EP-00301439.
PR	25-FEB-1999;	99US-0121835P.
PR	05-MAR-1999;	99US-012180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	21-APR-1999;	99US-0130077P.
PR	23-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	28-APR-1999;	99US-0130891P.
PR	30-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139113P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.

Query Match 73.6%; Score 1358; DB 3; Length 339;
 Best Local Similarity 85.5%; Pred. No. 5.5e-121;
 Matches 259; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

QY	3	SSSOHLVTTGGAGTCTHTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGGELLSONLQ 62
Db	4	SVEQNLVTTGGAGTCTHTVQVLLKDGFKVSIIDNFDNSVIEAVDRVRELVDGDLSSKLD 63
QY	63	FTQGLRNRDLEKLSKTTFFDAVIFHAGLKAFAESVAKPRRYFDFNLVGTINLYEFMAK 122
Db	64	FNLDLNRKGDIEKLFSKQRFDAVIFHAGLKAFAESVAKPRRYFDFNLVGTINLYETMAK 123
QY	123	YNCKMWFSSATVYQGEPIKCEBPKIQAMNPKYGRTKLFEEIARDIQKAEPEWKIL 182
Db	124	YNCKMWFSSATVYQGEPIKCEBPKIQAMNPKYGRTELFEEIARDIQKAEPEWRIVL 183
QY	183	LRYPNVPVGAHESKGLGDPKGIENNMPYIQQAVAGRLTFLNLYGHDPTRDGSARDYI 242
Db	184	LRYPNVPVGAHESGSGIDPKGIENNMPYIQQAVAGRLTFLNLYGHDPTRDGSARDYI 243
QY	243	HVMDLADGHTAALRKLTFTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
Db	244	HVMDLADGHTAALRKLTFTADPKICTAYNLGTGCTSVLEMTAFKASGKKIPVKLCPRR 303

```
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139450P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 05-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0143772P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144486P.
PR 18-JUL-1999; 99US-0144325P.
PR 18-JUL-1999; 99US-0144331P.
PR 18-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151030P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 06-OCT-1999; 99US-0157753P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.7%; Score 1231; DB 3; Length 375;
Best Local Similarity 66.4%; Pred. No. 8.9e-109;
Matches 229; Conservative 50; Mismatches 64; Indels 2; Gaps 2;
QY 2 VSSCHILVTGAGFICHTHTVOLLKAGFSVSIIDNEDNSWMEAVDVRQVVGPLLSONL 61
DB 26 VMARNVLVSGAGYIGSHTVQLLLGGYSVVVDNDSNSAVSLQRVKLAAB-HGERL 84
QY 62 QFTQDLENRDDLEKLFKSTTFDAVIHFAGLKAESVAKPRRYFDNLTGNTINLYEFMA 121
```

Db 85 SFHQVLDRLDRSALEKIFSETKFDVAIHFAGLKAVGESVEKPLLYNNLVGTTITLLEWMA 144
QY 122 KYNCKQWVFSSSATVYGOEKIPCEBDFKLQAMNPGVGRKLFLEBIARDIOKAEPEWKII 181
Db 145 QHCCKNLVFSSTATVYGSKEVPCTEEFISALNPGVGRKLFIEICRDVYGSDFEWKII 204
QY 182 LLRYFNPVGAHSGKLGEDPKGIPNNLMPIYIQVAVGRITELNLYGHDPTRDGSAIRDY 241
Db 205 LLRYFNPVGAHSGDITGEDPRGIPNNLMPIYVQVAVGRPHLTVEFGNDYNTKDGTVGRDY 264
QY 242 IHVMDLADGHIALRLKLTFTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPR 301
Db 265 IHVIDLADGHIALRLKLBCK-IGCEVYNLGTGNGTSVLEMTAFKASGKKIPLVIAGR 323
QY 302 RPDGATEVYASTERAKEILGWKANYGVEMCRDQWNAKNNPWGY 346
Db 324 RPDGAEVYVASTERASSELNWKAKYGIEMCRDLNWNASNNPYG 368

RESULT 7
AAG42840
ID AAG42840 standard; protein; 377 AA.
XX AC AAG42840;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53475.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145244P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-01444884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145376P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147360P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0155658P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.	QY	5	SOHLVTCGAGFCTHTVQLKAGPSVSIIDNFDSVMEAVDRVGVGPELLSONLOFT	64
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.	DB	3	ARNVLVGGAGYIGSHTVQLLLGGYSVVVDNLDSAVSLQKVKKLAAB-HGERLSFH	61
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.	QY	65	QGLDRNDDLEKLFSTTFDAVHFAGLKAVASVAKPRYFDFNLVGTINLYEFMAKYN	124
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.	DB	62	QVDLRDRSALEKIFSEIKFDVVIHFAGLKAVGESVEKPLLYNNLVGTITLLEVMAQHG	121
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.	QY	125	CKXWFFSSATVYQPEKICEEDFKLOANPYGRTKLFEEIARDIOKAEPWKILLR	184
PR	19-JUL-1999;	99US-0144336P.	PR	14-OCT-1999;	99US-0159331P.	DB	122	CKNLVFFSSATVYQSPKEVCTEFPISALNPYGRTKLFEEICRDVYGSDEPWKILLR	181
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.	QY	185	YFNPVGAHESGKLGEDPKGIPNNLMPIYIQVAVGRLTELNVYGHDPYTRDGSATRDYIHV	244
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159638P.	DB	182	YFNPVGAHPSODIGEDPRGIPNNLMPIYQVAVGRRPHLVFGNDYNTKDGTVRDYIHV	241
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159638P.	QY	245	MDLADGHIAALRKLTFTENIGCTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPG	304
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159638P.	DB	242	IDLADGHIAALRKLEDCCK-IGCEVYNLGTGNGTSVLEMTAFKASGKKIPLVIAGRRPG	300
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.	QY	305	DATEVYASTEAEBKELGWKANYGVVEEMCRDQWAKNPNWGY	346
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.	DB	301	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160815P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160980P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160981P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161361P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161920P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161922P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	05-AUG-1999;	99US-0147302P.	PR	28-OCT-1999;	99US-0161933P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	05-AUG-1999;	99US-0147192P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	05-AUG-1999;	99US-0147260P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	06-AUG-1999;	99US-0147303P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	06-AUG-1999;	99US-0147416P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	06-AUG-1999;	99US-0147493P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	08-AUG-1999;	99US-0147935P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	10-AUG-1999;	99US-0148171P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	11-AUG-1999;	99US-0148319P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	12-AUG-1999;	99US-0148341P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	13-AUG-1999;	99US-0148565P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	13-AUG-1999;	99US-0148684P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	16-AUG-1999;	99US-0149368P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	16-AUG-1999;	99US-0149175P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	18-AUG-1999;	99US-0149426P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	20-AUG-1999;	99US-0149722P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	20-AUG-1999;	99US-0149723P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	20-AUG-1999;	99US-0149829P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-AUG-1999;	99US-0149902P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-AUG-1999;	99US-0149930P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	25-AUG-1999;	99US-0150566P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	26-AUG-1999;	99US-0150884P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-AUG-1999;	99US-0151065P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-AUG-1999;	99US-0151066P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	27-AUG-1999;	99US-0151080P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	30-AUG-1999;	99US-0151303P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	31-AUG-1999;	99US-0151438P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	01-SEP-1999;	99US-0151930P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	07-SEP-1999;	99US-0152363P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	10-SEP-1999;	99US-0153070P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	13-SEP-1999;	99US-0153758P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	15-SEP-1999;	99US-0154018P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	16-SEP-1999;	99US-0154039P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	20-SEP-1999;	99US-0154779P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	22-SEP-1999;	99US-0155139P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	23-SEP-1999;	99US-0155486P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	24-SEP-1999;	99US-0155659P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	28-SEP-1999;	99US-0156458P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	29-SEP-1999;	99US-0156596P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	04-OCT-1999;	99US-0157117P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	05-OCT-1999;	99US-0157753P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	06-OCT-1999;	99US-0157865P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	07-OCT-1999;	99US-0158029P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	08-OCT-1999;	99US-0158232P.	PR	29-OCT-1999;	99US-0162142P.	DB	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342
PR	12-OCT-1999;	99US-0158369P.	PR	29-OCT-1999;	99US-0162142P.	QY	18	DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY	342

Query Match 66.6%; Score 1229; DB 3; Length 349;

Best Local Similarity 66.7%; Pred. No. 1.2e-108;

Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;

QY 5 SOHLVTCGAGFCTHTVQLKAGPSVSIIDNFDSVMEAVDRVGVGPELLSONLOFT 64

DB 3 ARNVLSGGAGYIGSHTVQLLLGGYSVVVDNLDSAVSLQKVKKLAAB-HGERLSFH 61

QY 65 QGLDRNDDLEKLFSTTFDAVHFAGLKAVASVAKPRYFDFNLVGTINLYEFMAKYN 124

DB 62 QVDLRDRSALEKIFSEIKFDVVIHFAGLKAVGESVEKPLLYNNLVGTITLLEVMAQHG 121

QY 125 CKXWFFSSATVYQPEKICEEDFKLOANPYGRTKLFEEIARDIOKAEPWKILLR 184

DB 122 CKNLVFFSSATVYQSPKEVCTEFPISALNPYGRTKLFEEICRDVYGSDEPWKILLR 181

QY 185 YFNPVGAHESGKLGEDPKGIPNNLMPIYIQVAVGRLTELNVYGHDPYTRDGSATRDYIHV 244

DB 182 YFNPVGAHPSODIGEDPRGIPNNLMPIYQVAVGRRPHLVFGNDYNTKDGTVRDYIHV 241

QY 245 MDLADGHIAALRKLTFTENIGCTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPG 304

DB 242 IDLADGHIAALRKLEDCCK-IGCEVYNLGTGNGTSVLEMTAFKASGKKIPLVIAGRRPG 300

QY 305 DATEVYASTEAEBKELGWKANYGVVEEMCRDQWAKNPNWGY 346

DB 301 DAEVYIASTEAEBELNWKANYGIEEMCRDLNWKASNNPYGY 342

RESULT 10

AAG42842

ID AAG42842 standard; protein; 350 AA.

XX AAG42842;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 53477.

XX Protein identification; signal transduction pathway; metabolic pathway;

XX hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX EPI033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139753P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141272P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147415P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.

PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	28-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0161040P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161932P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 66.4%; Score 1229; DB 3; Length 350;			
Best Local Similarity 66.7%; Pred. No. 1.2e-108;			
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;			
Qy	5	SOHLVTGGAGFTGTHVWOLLKAGFSVSIIDNFDNSVMEANDRVQVVGPLLSQLQFT	64
Db	2	ARNVLVGGAGYTGSHTVLQLLGGYSVVVDNLDNSAVSQRVKKLAAE-HGERLSPH	60
Qy	65	QGLRNRDDLEKLFSTTFDAVHFAGLKAVASVAKPRRYFDNVLVGTINLYEFNAKYN	124
Db	61	QVDLRDRSALEKTFSETKFDVHFAGLKAVGESVEKPLLYNNLVGNTITLLEVMAQHG	120
Qy	125	CKMVFSSSTVTVGQPEKPCBEDDFKQANVPYGRKLFLEETARDIQKAEPEWKILLR	184
Db	121	CKNLVFSSATVYGSPEVCTEFPISALNPYGRKLFEEICRDVYGSDEPEWKILLR	180
Qy	185	YFNPVGAHESGKLGDEPKGTPNNLMPYIQOVAVGRLTETNLVYGHDPYTRDGSAIRDIHV	244
Db	181	YFNPVGAHPGSDIGEDRGRIPNNLMPFVQOVAVGRPHLTVFGNDYTKDGTGRDIHV	240
Qy	245	MDLADGHIAALRKLFTEENTGCTAYNLGTGRGTSVLEMTVAFKAGKKIPVKLCPRRPG	304
Db	241	IDLADGHIAALRKLEDCCK-IGCEVYNLTGNGTSVLEWDAFEKASGKKIPLVIAGRRPG	299
Qy	305	DATEVYASTERAELKELGKANYGVEMCRDWNWAKNNPWGY	346
Db	300	DAEVYASTERAELSNLWAKKAGIEMCRDLNWNWASNNPYGY	341

ID	AAG42841 standard; protein; 351 AA.
XX	AAG42841;
AC	
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 53476.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127482P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	04-MAY-1999; 99US-0132407P.
PR	05-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137282P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.
PR	18-JUN-1999; 99US-0139459P.
PR	18-JUN-1999; 99US-0139460P.

PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140333P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140334P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140931P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142330P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143644P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144335P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145941P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
QY	5	SOHILVTGGAGFIGTHTVYQLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQNLQFT	QY	5	SOHILVTGGAGFIGTHTVYQLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSQNLQFT
Db	3	ARNVLVSGAGYIGSHTVQLLLGGYSVVVDNLDNDSAVSLQVRVKLAAB-HGERLSFH	Db	3	ARNVLVSGAGYIGSHTVQLLLGGYSVVVDNLDNDSAVSLQVRVKLAAB-HGERLSFH
QY	65	QGDLRNRDLEKLPKTTTFDAVIHPAGLKAVAESVAKPRRYPDFNLVGTINLYEFMAKN	QY	65	QGDLRNRDLEKLPKTTTFDAVIHPAGLKAVAESVAKPRRYPDFNLVGTINLYEFMAKN
Db	62	QVDLRDRSALEKIFSETKFDVAIHPAGLKAVGESVEKPLLYNNLVGTITLLEYVAQHG	Db	62	QVDLRDRSALEKIFSETKFDVAIHPAGLKAVGESVEKPLLYNNLVGTITLLEYVAQHG
QY	125	CKKMFSSSATVYGOPEKIPCEEDFKLOAMPYGTGKLFLEBIARDIOKAEPEWKIILLR	QY	125	CKKMFSSSATVYGOPEKIPCEEDFKLOAMPYGTGKLFLEBIARDIOKAEPEWKIILLR
Db	122	CKNLFSSSATVYGSKEVPCTEEPISALAPYGTGKLFIEICRDVYGSPEWKIILLR	Db	122	CKNLFSSSATVYGSKEVPCTEEPISALAPYGTGKLFIEICRDVYGSPEWKIILLR
QY	185	YFNPVGAHESGKLGEDPKGIPNNLMPYIQQVAVGSLTELNAVYGHDPTRDGSARDYTHV	QY	185	YFNPVGAHESGKLGEDPKGIPNNLMPYIQQVAVGSLTELNAVYGHDPTRDGSARDYTHV

Query Match 66.6%; Score 1229; DB 3; Length 351;

Best Local Similarity 66.7%; Pred. No. 1.3e-108;

Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;

Db 182 YFNPVGAHPSGDIGEDPRGIPNNLMPFVQOVAVGRPHLTVPFGDNTKDGTVGRDYIHV 241
Qy 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRRP 304
Db 242 IDLADGHIAALRKLDECK-IGEVNLTGNGTSVLENDVDFKASGKKIPLVIAGRRP 300
Qy 305 DATEVYASTERAEKELGWKANGYVEEMCRDQWNWAKNNPWGY 346
Db 301 DAEVYVYASTERAESELNWKAKYGIEMCRDLNWNASNNPYG 342

RESULT 12

AAG53259
ID AAG53259 standard; protein; 348 AA.

XX AC AAG53259;

XX DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 67789.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 28-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134321P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 21-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

Query Match	Score	1222;	DB 3;	Length	348;
Best Local Similarity	65.4%;				
Matches	225;	Conservative	51;	Mismatches	66;
				Indels	2;
				Gaps	2;
7 HITVTCGAGTGTHTVVLKAGFSVSIIDNFNSVMEAVDRVGVGPLLSONLOFTOG	66.2%;				
99US-0147935P.					
99US-0148171P.					
99US-0148319P.					
99US-0148341P.					
99US-0148565P.					
99US-0148684P.					
99US-0149368P.					
99US-0149175P.					
99US-0149426P.					
99US-0149722P.					
99US-0149723P.					
99US-0149902P.					
99US-0149929P.					
99US-0149920P.					
99US-0149930P.					
99US-0150566P.					
99US-0150884P.					
99US-0151065P.					
99US-0151066P.					
99US-0151080P.					
99US-0151303P.					
99US-0151438P.					
99US-0151930P.					
99US-0152363P.					
99US-0153070P.					
99US-0153758P.					
99US-0154018P.					
99US-0154039P.					
99US-0154779P.					
99US-0155139P.					
99US-0155486P.					
99US-0155659P.					
99US-0156458P.					
99US-0156596P.					
99US-0157117P.					
99US-0157753P.					
99US-0157865P.					
99US-0158029P.					
99US-0158232P.					
99US-0158369P.					
99US-0159293P.					
99US-0159294P.					
99US-0159295P.					
99US-0159329P.					
99US-0159330P.					
99US-0159637P.					
99US-0159638P.					
99US-0159584P.					
99US-0160741P.					
99US-0160767P.					
99US-0160768P.					
99US-0160770P.					
99US-0160814P.					
99US-0160815P.					
99US-0160980P.					
99US-0160981P.					
99US-0160989P.					
99US-0161404P.					
99US-0161405P.					
99US-0161406P.					
99US-0161559P.					
99US-0161360P.					
99US-0161361P.					
99US-0161320P.					
99US-0161922P.					
99US-0161933P.					
99US-0162142P.					
99US-0162143P.					
99US-0162144P.					
99US-0162145P.					
99US-0162146P.					
99US-0162147P.					
99US-0162148P.					
99US-0162149P.					
99US-0162150P.					
99US-016					

PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140659P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148119P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.

PR 05-MAR-1999; 99US-01231802P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139432P.
PR 18-JUN-1999; 99US-0139434P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145278P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147418P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149262P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154773P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159337P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity 66.1%; Score 1220; DB 3; Length 372;

Matches 225; Conservative 51; Mismatches 66; Indels 2; Gaps 2;

Qy	7	HILVTGGAGTGTHTVVQLLKAGFSVSIIDNPNVSVMKAVDRVQVVGPLLSQNLQFTQG 66
Db	28	NILVTGGAGTGTHTVVQLLKAGFSVSIIDNPNVSVMKAVDRVQVVGPLLSQNLQFTQG 66
Qy	67	DLNRDDLEKLFKTTFDVIFHAGLKAVASVAKPRRYDFNLTGNTINLYEFMAKYNCK 126
Db	87	DLRDKPALEKVFSETKFDVIMHAGLKAVGESVAKPLLYNNNLATITLLEVMAHCK 146
Qy	127	KWVFSSATVYGOEKPCEDEKIQAMNPKYGTCLFLZEIARDIQKAEPEWKIILLRYF 186
Db	147	KLVFSSATVYGPKEVCFTEESFLSGMSPYGTCLFIEDICRDVORGDPWRIIMLRYF 206
Qy	187	NPVGAHESGKLGEDPKGIPNNLMPYIQQVAVGRLTELVYGHDPYPRDGSAIRDYIHVD 246
Db	207	NPVGAHPSGRIGEDPCTPNNLMPYVQVVGRLPNLKIYGTDTTKDGTGVDRYIHVD 266
Qy	247	LADGHIAALAKLFTTENIGTATNLTGRTSVLENMTAFKASGKKIPVKLCPRRPGDA 306
Db	267	LADGHIFALQKLDDE-IGCEVNLGTGKTTVLEWDAFEKASGMKIPLVKVGRRPGDA 325
Qy	307	TEVYASTERAELKGNKANYGVEEMCRDQNNWAKNPNWGYAGKP 350
Db	326	ETVYASTERAELNWKANFGIEEMCRDQNNWASNNPFGYGSSP 369

Search completed: March 22, 2004, 21:53:48
Job time : 78 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:52:35 ; Search time 23 Seconds
(without alignments)
785.613 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSHILVTGGAGFIGTH.....MCRDOWNAKNNPMWYAGRP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	977.5	53.0	348	3	US-09-113-536-2
2	977.5	53.0	348	4	US-09-624-183-2
3	977.5	53.0	348	5	PCT-US95-05785-2
4	945	51.2	376	4	US-09-540-236-2180
5	926.5	50.2	345	4	US-09-543-681A-4948
6	918	49.8	363	4	US-09-489-039A-13195
7	893.5	48.4	338	4	US-09-580-929-4
8	833	34.3	469	4	US-09-107-532A-5936
9	624	33.8	341	4	US-09-134-000C-5524
10	619	33.6	329	4	US-09-134-000C-6238
11	553.5	30.0	359	4	US-09-634-955B-21
12	449.5	24.4	341	4	US-09-328-352-7330
13	334	18.1	329	4	US-09-107-532A-7038
14	318	17.2	344	4	US-09-489-039A-8087
15	311	16.9	99	4	US-09-424-311-3
16	308.5	16.7	347	4	US-09-543-681A-4531
17	306	16.6	339	4	US-09-489-039A-8139
18	301.5	16.3	339	4	US-09-252-991A-26600
19	278	15.1	341	4	US-09-580-929-3
20	278	15.1	352	4	US-09-580-929-2
21	272	14.7	290	4	US-09-134-000C-4233
22	268.5	14.6	91	4	US-09-424-311-4
23	266.5	14.4	276	4	US-09-134-000C-3713
24	262	14.2	346	4	US-09-328-352-5625
25	244	13.2	370	4	US-09-543-681A-4353
26	218.5	11.8	329	3	US-09-036-987A-26
27	218.5	11.8	329	3	US-09-370-700-26

Sequence 26, Appli
Sequence 23322, A
Sequence 19041, A
Sequence 5, Appli
Sequence 8053, Ap
Sequence 4, Appli
Sequence 6233, Ap
Sequence 15, Appl
Sequence 17, Appl
Sequence 15, Appl
Sequence 14, Appl
Sequence 35, Appl
Sequence 11, Appl
Sequence 10678, A
Sequence 7414, Ap
Sequence 7549, Ap
Sequence 17956, A
Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-113-536-2

; Sequence 2, Application US/09113536

; Patent No. 6153739

; GENERAL INFORMATION:

; APPLICANT: JI, H., ET AL.

; TITLE OF INVENTION: HUMAN UDP GALACTOSE-4-EPIMERASE

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/113.536

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/462.966

; FILING DATE: June 5, 1995

; APPLICATION NUMBER: PCT/US95/05785

; FILING DATE: 11 MAY 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: FERRARO, GREGORY D.

; REGISTRATION NUMBER: 36,134

; REFERENCE/DOCKET NUMBER: 325800-430

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 348 AMINO ACIDS

; TYPE: AMINO ACID

; STRANDEDNESS:

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

US-09-113-536-2

Query Match 53.0%; Score 977.5; DB 3; Length 348;
Best Local Similarity 53.9%; Pred. NO. 4.2e-91;
Matches 188; Conservative 58; Mismatches 90; Indels 13; Gaps 4;

QY 124 -----NCKMVFSSATVYQGPKEIPCE---EDFKLQANMPYGRTKLFLBEIARDIQ 172
Db 119 SALDETAKGVKVFSSDEVYGDLESIPISAFTEDPYNPSSPYGASKASSELVYRAYH 178
QY 173 KAPEPKIILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQVAVGRLTETELNYGHDYPT 232
Db 179 RAY-GLPAILIRYFNYGYSGRIGEDNGEPKILPILIONALKGEPLVPYGDYPT 237
QY 233 RDGSAIRDYHVMDLA-DGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASG 291
Db 238 PDGTQVRDTHVEDHARANHLALT-K-GRACKGSEVYNNIGGNEYSNLEVEATEKLLG 295
QY 292 -----KKIPVKLCPRPGDATEVYVASTERAELGWKANY-GVEEMCRDQWNA 339
Db 296 ELAPKPHVKAKEDPATFVDDRPDDARYAADASKIKRELGWKPEVTNLEEGLATVNYWY 355
QY 340 KNN 342
Db 356 LEN 358

RESULT 12

US-09-328-352-7330
; Sequence 7330, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7330
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7330

Query Match 24.4%; Score 449.5; DB 4; Length 341;
Best Local Similarity 31.2%; Pred. No. 2.5e-37;
Matches 105; Conservative 74; Mismatches 147; Indels 11; Gaps 5;
QY 8 ILVTGAGIGTHTVOLLKAGFSVSIIDFNFSVMEAVDRVRQVVGPLLSONLOFTQGD 67
Db 8 ILVTGGLGFIGSHIALSLMAQGVVLDNLANSTLQTLERLEYISG---MYVPFVKLD 63
QY 68 LNRDDLEKLFKSTTFDAVIHFAGLKAVASVAKPRRYFDNLVGTINLYEFMAKYNCK 127
Db 64 VENTPALNVFQYSDAIVIHAGFKSIEESNLKPLEYNDVNSCINLLRAMQRTGVH 123
QY 128 MVFSSATVYQGP-EKIPCEEDFKLQANMPYGRTKLFLBEIARDIQKABPEWKILLRYF 186
Db 124 FHLSSLAVYKSGSLSQSETDFNAYPNFYTKSQQMIEIIRDTYKIDHEWKIAILRLS 183
QY 187 NPVGASHGKLGEDPKGIPNNLMPYIQVAVR--LTELNVYGHDPYPTDGSARDYHV 244
Db 184 NIVGAFEHVGLGYAQLPKNLIVPLAMQVAAQRDLIELQ---DOSSTEDHTFVSFLHV 240
QY 245 MDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRPG 304
Db 241 LDVCEAVSSTLHWL-RDQSHCCBAFNAHEQVHSIRQLLDEISQVTOAEIPTQSAIYKHE 299
QY 305 DATEVYASTERAELGWKANYGVEEMCRDQWNAK 341
Db 300 ELAQIGANTGKATLLQWVFKRTLQMIEDERFYQN 336

RESULT 13

US-09-107-532A-7038
; Sequence 7038, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7038:
SEQUENCE CHARACTERISTICS:
LENGTH: 329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...329
SEQUENCE DESCRIPTION: SEQ ID NO: 7038:
US-09-107-532A-7038
Query Match 18.1%; Score 334; DB 4; Length 329;
Best Local Similarity 27.1%; Pred. No. 1.4e-25;
Matches 90; Conservative 72; Mismatches 124; Indels 46; Gaps 10;
QY 9 LVTGGAGFIGTHTVOLLKAGFSVSIIDFNFSVMEAVDRVRQVVGPLLSONLOFTQGD 68
Db 19 LITGGAGFIGS-TLANYLGKNNVVVVDLSMGKKENLDE-----SKHITFIEGDV 68
QY 69 RNRDDLEKLFKSTTFDAVIHFAGLKAVASVAKPRRYFDNLVGTINLYEFMAKYNCK 126
Db 69 ADKNLEKVMKEVHFYIPLAAVASVADSVVERPLETHRVNFPDALLLELVKYSGLK 128
QY 127 KMFSSATVYQGP-EKIPCEEDFKLQANMPYGRTKLFLBEIARDIQKABPEWKILLRYF 186
Db 129 RLVFSSAAVYGDPTLPKKEESVIRPLTPYADKFAAQVYVLDYCHLY-DVPTSAVRFF 187
QY 187 NPVGASHGKLGEDPKGIPNNLMPYIQVAV-----GRLTENVYGHDPYPTDGS 236
Db 188 NVTGPNQN-----PNS--PYGVSILVDRYKKQLAGEKTEFTIFG-----DGS 229
QY 237 AIRDYIHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPV 296
Db 230 QSRDFVYIEDV---TQALLVAKKEKALCQQQNVGTGKSTLLELIHSDIQLGTAL 285
QY 297 KLCFRPGDATEVYVASTERAELGWKANYGV 328
Db 286 KYAERSGDIRDLSADISKI-RSLGYQPKPDI 316

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:55:36 ; Search time 63 Seconds
(without alignment)
1438.641 Million cell updates/sec

Title: US-09-913-064A-14
Perfect score: 1845
Sequence: 1 MVSSSQHLVTGGAGFIGTH.....MCRQWNWAKNPNWYACKP 350

Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1840	99.7	350	12	US-10-424-599-179426
2	1762	95.5	363	12	US-10-425-114-47644
3	1486	80.5	351	14	US-10-060-275-2
4	1198.5	65.0	334	12	US-10-425-114-46573
5	1196.5	64.9	334	12	US-10-425-114-39413
6	1192	64.6	391	12	US-10-424-599-232009
7	1192	64.6	408	12	US-10-425-114-55605
8	1190	64.5	408	12	US-10-425-114-56199
9	1183.5	64.1	441	12	US-10-424-599-232008
10	1148	62.2	403	12	US-10-425-114-59200
11	1141	61.8	332	12	US-10-425-114-61106
12	1071.5	58.1	333	12	US-10-425-114-58229
13	1040	56.4	326	12	US-10-425-114-41119
14	986.5	53.5	348	12	US-09-318-271-10
15	986.5	53.5	348	14	US-10-303-664A-18

16	977.5	53.0	348	13	US-10-201-298-2	Sequence 2, Appli
17	918.5	49.8	350	13	US-10-108-605-65	Sequence 65, Appl
18	893.5	48.4	338	9	US-09-318-271-8	Sequence 8, Appli
19	887.5	48.1	336	15	US-10-429-812-10	Sequence 10, Appl
20	878.5	47.6	699	15	US-10-369-493-1430	Sequence 1430, Ap
21	836.5	45.3	255	12	US-10-424-599-148894	Sequence 148894,
22	775.5	42.0	207	12	US-10-424-599-232005	Sequence 232005,
23	775.5	42.0	207	12	US-10-425-114-45766	Sequence 45766, A
24	761	41.2	228	12	US-10-424-599-232010	Sequence 232010,
25	751	40.7	286	12	US-10-425-114-55233	Sequence 55233, A
26	683.5	37.0	216	12	US-10-425-114-44206	Sequence 44206, A
27	645.5	35.0	170	12	US-10-425-114-43920	Sequence 43920, A
28	634	34.4	415	12	US-10-424-599-273943	Sequence 273943,
29	625.5	33.9	513	12	US-10-425-114-62570	Sequence 62570, A
30	616.5	33.4	497	12	US-10-425-114-67924	Sequence 67924, A
31	615	33.3	327	15	US-10-369-493-14762	Sequence 14762, A
32	614	33.3	209	12	US-10-425-114-47486	Sequence 47486, A
33	612.5	33.2	438	12	US-10-425-114-60686	Sequence 60686, A
34	608	33.0	327	15	US-10-369-493-16982	Sequence 16982, A
35	569	30.8	325	15	US-10-369-493-19845	Sequence 19845, A
36	562.5	30.5	319	14	US-10-156-761-11112	Sequence 11112, A
37	557	30.2	423	12	US-10-424-599-273945	Sequence 273945,
38	553.5	30.0	359	14	US-10-172-585-21	Sequence 21, Appl
39	544.5	29.5	329	9	US-09-738-626-5610	Sequence 5610, Ap
40	534.5	29.0	325	15	US-10-369-493-8085	Sequence 8085, Ap
41	530.5	28.8	326	14	US-10-156-761-9201	Sequence 9201, Ap
42	530.5	28.5	321	15	US-10-369-493-10165	Sequence 10165, A
43	501.5	27.2	326	14	US-10-156-761-13620	Sequence 12620, A
44	487.5	27.0	258	12	US-10-425-114-63671	Sequence 63671, A
45	491.5	26.6	317	15	US-10-369-493-21857	Sequence 21857, A

ALIGNMENTS

RESULT 1
US-10-424-599-179426
; Sequence 179426, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 179426
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_133037C.1.pep
US-10-424-599-179426

Query Match	99.7%	Score	1840	DB	12	Length	350
Best Local Similarity	99.7%	Pred. No.	8.2e-187				
Matches	349	Conservative	0	Mismatches	1	Indels	0
Qy	1	MVSSSQHLVTGGAGFIGTHVTVQLLKAGFSVSIIDFNSVMEAVDRVQVVGFLLSQN	60				
Db	1	MVSSSQHLVTGGAGFIGTHVTVQLLKAGFSVSIIDFNSVMEAVDRVQVVGFLLSQN	60				
Qy	61	LQFTQGLNRDLDLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDFNLVTGTLNLFYM	120				
Db	61	LQFTQGLNRDLDLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDFNLVTGTLNLFYM	120				
Qy	121	AKYNCKKXVSSSATVVGQPEKIPCEBDFKLOANPNYGRTKLFLEETARDIQKAEPSWKI	180				
Db	121	AKYNCKKXVSSSATVVGQPEKIPCEBDFKLOANPNYGRTKLFLEETARDIQKAEPSWKI	180				

QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 240
DB 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 240
QY 241 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
DB 241 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
QY 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 350
DB 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 350

RESULT 2

US-10-425-114-47644
; Sequence 47644, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47644
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700787631_FLI.pgp
US-10-425-114-47644

Query Match 95.5%; Score 1762; DB 12; Length 363;
Best Local Similarity 95.1%; Pred. No. 1.7e-178;
Matches 333; Conservative 9; Mismatches 8; Indels 0; Gaps 0;

QY 1 MYSSQHILVTGGAGFIGTHTVVQLLKAGFSVSIIDNFNSVMEAVDRVRQVGVPLLSQN 60
DB 14 WTSSQKILVTGGAGFIGTHTVVQLLKAGFSVSIIDNFNSVMEAVDRVRQVGVPLLSQN 73
QY 61 LFTQGLRNRDDLEKLFSTTFDAVTHPAGLKAVAESVAKPRRYDFNLVGTINLYEFM 120
DB 74 LFTQGLRNRDDLEKLFSTTFDAVTHPAGLKAVAESVAKPRRYDFNLVGTINLYQVM 133
QY 121 AKYCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 180
DB 134 AKYCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 193
QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 240
DB 194 ILLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 253
QY 241 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 300
DB 254 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 313
QY 301 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 350
DB 314 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 363

RESULT 3

US-10-060-275-2
; Sequence 2, Application US/10060275
; Publication No. US20030073828A1
; GENERAL INFORMATION:

; APPLICANT: DAI, ZIYU
; APPLICANT: SHI, LIFANG
; APPLICANT: HOOKER, BRIAN S.
; TITLE OF INVENTION: EPIMERASE GENE AND USE THEREOF
; FILE REFERENCE: 059440-0143
; CURRENT APPLICATION NUMBER: US/10/060,275
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 60/265,311
; PRIOR FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-10-060-275-2

Query Match 80.5%; Score 1486; DB 14; Length 351;
Best Local Similarity 78.2%; Pred. No. 4e-149;
Matches 273; Conservative 33; Mismatches 43; Indels 0; Gaps 0;

QY 2 VSSSQHILVTGGAGFIGTHTVVQLLKAGFSVSIIDNFNSVMEAVDRVRQVGVPLLSQN 61
DB 3 VQCOENILVTGGAGFIGTHTVVQLLNKGFVSIIDNFNSVMEAVDRVRQVGVPLLSQN 62
QY 62 QFTQGLRNRDDLEKLFSTTFDAVTHPAGLKAVAESVAKPRRYDFNLVGTINLYEFMA 121
DB 63 EFLGDIRNRDDLEKLFSTTFDAVTHPAGLKAVGESVVPFLYFENNLGSLTLYSVMA 122
QY 122 KYNCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 181
DB 123 KYNCKKRVFSSSATVYGOPEKIPCEEDFKLQANPYGRTKLFLEIARDIQKAPENKI 182
QY 182 LLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 241
DB 183 LLRYFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVGHDPTRDGSARD 242
QY 242 IHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 301
DB 243 IHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCP 302
QY 302 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 350
DB 303 RRGDATEVYASTERAELGKWKANYGVVECMCRDQWNWAKNPNWGYAGKP 351

RESULT 4

US-10-425-114-46573
; Sequence 46573, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46573
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700445105_FLI.pgp
US-10-425-114-46573

Query Match 65.0%; Score 1198.5; DB 12; Length 394;
Best Local Similarity 66.0%; Pred. No. 1.9e-118;

Matches 231; Conservative 44; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 119
Db 99 NLVFKVLDLRDLRALVDIFSSHFRAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLV 298
Db 279 DYHWMDLADGHIAALRKLYEDSKIGCEVYNLGTGRGTSVLEMTAFKASGKKIPLV 338
QY 299 CRRPGDAEVYASTRAEKELGWKANYGVEMCRDQWNAKKNPWGYAG 348
Db 339 AGRPGDAEIVYATAKAEKELGWKANYGVEMCRDQWNAKKNPWGYAG 388
RESULT 5
US-10-425-114-39413
; Sequence 39413, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39413
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI.pep
US-10-425-114-39413
Query Match 64.9%; Score 1196.5; DB 12; Length 394;
Best Local Similarity 65.7%; Pred. No. 3.1e-118;
Matches 230; Conservative 45; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 119
Db 99 NLVFKVLDLRDLRALVDIFSSHFRAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLV 298

Db 279 DYHWMDLADGHIAALRKLYEDSKIGCEVYNLGTGRGTSVLEMTAFKASGKKIPLV 338
QY 299 CRRPGDAEVYASTRAEKELGWKANYGVEMCRDQWNAKKNPWGYAG 348
Db 339 AGRPGDAEIVYATAKAEKELGWKANYGVEMCRDQWNAKKNPWGYAG 388
RESULT 6
US-10-424-599-232009
; Sequence 232009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 232009
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51526C.1.pap
US-10-424-599-232009
Query Match 64.6%; Score 1192; DB 12; Length 391;
Best Local Similarity 66.4%; Pred. No. 9.2e-118;
Matches 225; Conservative
QY 8 ILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQNLQFTQGD 67
Db 48 VLVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQNLQFTQGD 106
QY 68 LRNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEFMAKYNCK 127
Db 107 LRNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEFMAKYNCK 166
QY 128 MVSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWKIILLYEFN 187
Db 167 LVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWKIILLYEFN 226
QY 188 PVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIRDIHYMDL 247
Db 227 PVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIRDIHYMDL 286
QY 248 ADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVLCRRPGDAT 307
Db 287 ADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLVLCRRPGDAT 345
QY 308 EYASTRAEKELGWKANYGVEMCRDQWNAKKNPWGY 346
Db 346 EYASTRAEKELGWKANYGVEMCRDQWNAKKNPWGY 384
RESULT 7
US-10-425-114-55605
; Sequence 55605, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55605
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700103191_FLI.pep
US-10-425-114-55605
Query Match 64.9%; Score 1196.5; DB 12; Length 394;
Best Local Similarity 65.7%; Pred. No. 3.1e-118;
Matches 230; Conservative 45; Mismatches 72; Indels 3; Gaps 3;
QY 1 MVSS-SQHILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 59
Db 40 MVSALVLTILVTGGAGTGTHTVQVOLLKAGSVSIIDNFDNSVMEAVDRVQVGGPLLSQ 98
QY 60 NLOFTQGLNRDLEKLFSTTFDAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 119
Db 99 NLVFKVLDLRDLRALVDIFSSHFRAVTHFAGLKAVASVAKPRRYDFNVLGTINLYEF 158
QY 120 MAKYNCKMVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 179
Db 159 MAANGCKLVFSSATVYGPEKIPCEEDFKLOANPYGRTKLFLBETARDIOKAEPEWK 218
QY 180 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 239
Db 219 IILLYRFPVGAHSGKLGEDPKGIPNNLMPYIQOVAVGRLTFLNLYGHDYPTDRGSAIR 278
QY 240 DYHWMDLADGHIAALRKLF-TTENIGCTAYNLGTGRGTSVLEMTAFKASGKKIPLV 298

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:44:15 ; Search time 27 Seconds
(without alignments)
1246.927 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MYSSQHILVTGGAGFIGTH.....MCRDQWNAKNPWGAGKP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1671	90.6	350	2 T06526	UDPglucose 4-epime
2	1578	85.5	351	2 B86261	UDPglucose 4-epime
3	1571	85.1	351	2 S62783	UDPglucose 4-epime
4	1522	82.5	353	2 D96657	hypothetical prote
5	1500	81.3	354	2 T10426	UDPglucose 4-epime
6	1246	67.5	350	2 T10428	UDPglucose 4-epime
7	1229	66.6	350	2 T01881	UDPglucose 4-epime
8	1229	66.6	351	2 T04291	probable UDPglucos
9	1208	65.3	339	2 T08911	UDPglucose 4-epime
10	1038	56.3	339	2 D96628	UDPglucose 4-epime
11	983	53.3	338	2 F82419	UDP-glucose 4-epim
12	971.5	52.7	335	2 AC3357	UDPglucose 4-epime
13	966	52.4	339	2 B35187	UDPglucose 4-epim
14	964	52.2	339	2 C98034	UDPglucose 4-epime
15	962	52.1	347	2 T19989	hypothetical prote
16	943	51.1	347	2 S11223	UDPglucose 4-epime
17	941	51.0	336	2 D95213	UDPglucose 4-epim
18	928.5	50.3	339	2 F82014	UDPglucose 4-epime
19	920.5	49.9	338	1 A64063	UDPglucose 4-epim
20	916.5	49.7	339	1 S39638	UDPglucose 4-epim
21	915.5	49.7	339	2 S42430	UDPglucose 4-epim
22	910.5	49.3	337	1 A36931	UDPglucose 4-epime
23	910.5	49.3	338	2 AF0594	UDPglucose 4-epime
24	908	49.2	326	2 E98077	UDPglucose 4-epime
25	904.5	49.0	338	1 S51328	UDPglucose 4-epime
26	904.5	49.0	338	1 S34984	UDPglucose 4-epime
27	894.5	48.5	336	1 S70744	UDPglucose 4-epime
28	894.5	48.5	338	2 C90727	UDP-galactose-4-ep
29	894.5	48.5	338	2 D85578	UDP-galactose-4-ep

RESULT 2

ALIGNMENTS

RESULT 1

T06526

UDPglucose 4-epimerase (EC 5.1.3.2) - garden pea

N/Alternate names: UDPgalactose 4-epimerase

C/Species: Pisum sativum (Garden pea)

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000

C/Accession: T06526

R/Slocum, R.D.; Lake, M.R.; Williamson, C.L.

submitted to the EMBL Data Library, July 1995

A/Description: Cloning and characterization of a UDP-galactose-4-epimerase ('galactowal'

A/Reference number: Z15735

A/Accession: T06526

A/Status: preliminary; translated from GE/EMBL/DDBJ

A/Molecule type: mRNA

A/Residues: 1-350 <SLO>

A/Cross-references: EMBL:U31544; NID:g1173554; PIDN:AAA86532.1; PID:g1173555

C/Genetics:

A/Gene: galE

C/Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

C/Keywords: galactose metabolism; isomerase

F/8-346/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 90.6%; Score 1671; DB 2; Length 350;

Best Local Similarity 90.0%; Pred. No. 2e-129;

Matches 315; Conservative 16; Mismatches 19; Indels 0; Gaps 0;

QY 1 MYSSQHILVTGGAGFIGTHVTVOLLKAGFSVSIIDFNDSVMEAVDRVRYVGPILLSON 60

Db 1 MVASSQKILVTGSAGFTGTHVTVOLLNNGFNVSIIIDFNDSVMEAVDRVRYVGPILLSON 60

QY 61 LQFTQGLNRDDEKLFSTTDFDAVHFAGLKAVESVAKPRRYDFNLVGTINLYEPM 120

Db 61 LEFTGLDRLNKDDLEKLFSTKSDAVHFAGLKAVGESVENPRRYFDNLLVGTINLYEVM 120

QY 121 AKYNCKWVSSSNVTYVQPEKIPCEDEKLOANNPYGRTKLFLEIARDIOKAEPEWKI 180

Db 121 AKRNCKWVSSSNVTYVQPEKIPCEDEKLOANNPYGRTKLFLEIARDIOKAEPEWKI 180

QY 181 ILLRYFPVGAHESGKLGEDPKGI PNNLMFYIQQAVAGRLTINLVYGHDPYTRDGSARD 240

Db 181 VLLRYFPVGAHESGKLGEDPRGI PNNLMFYIQQAVAGRLPELVYGHDPYTRDGSARD 240

QY 241 YIHYMDLADGHIAALRKLFTEENIGCTAVNLGRTGTSVLEVTAPKASGKKIPVKLC 300

Db 241 YIHYMDLADGHIAALRKLFTEENIGCTAVNLGRTGTSVLEVTAPKASGKKIPVKLC 300

QY 301 RRPQDATEVYASTERAKELGWKNKANYGVBEEMCRDQWNAKNPNWYAGKRP 350

Db 301 RRPQDATEVYASTERAKELGWKNKANYGVBEEMCRDQWNAKNPNWYAGKRP 350

B86261
UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: B86261
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <STO>
A:Cross-references: GB:AB005172; NID:98698725; PIDN:AAF78483.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
A:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: isomerase

Query Match 85.5%; Score 1578; DB 2; Length 351;
Best Local Similarity 85.6%; Pred. No. 8.7e-122;
Matches 297; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

QY 3 SSSOHILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 62
DB 4 SVEQNILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 63
QY 63 FTQGLNRDLEKLSKTTFDVAVHPAGLKAVASVAKPRRYFDNVLGTTINLYETMAK 122
DB 64 FNLGDLRNKGDIEKLFKQDFDAVHPAGLKAVGESVEKGRYFDNVLGTTINLYETMAK 123
QY 123 YNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKIL 182
DB 124 YNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKIL 183
QY 183 LRYFNPVGAHESKGLGDPKGI PNNLMPYIQQVAVGRLPELVYGHDPYTRDGSARDYI 242
DB 184 LRYFNPVGAHESKGLGDPKGI PNNLMPYIQQVAVGRLPELVYGHDPYTRDGSARDYI 243
QY 243 HVMDLADGHTAALRKLFTTENGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
DB 244 HVMDLADGHTAALRKLFTTENGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 303
QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 349
DB 304 SGDATAVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 350

RESULT 3
S62783
UDPglucose 4-epimerase (EC 5.1.3.2) - Arabidopsis thaliana
N:Alternate names: uridine diphosphate glucose epimerase
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Aug-1996 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C:Accession: S62783
R:Doermann, P.; Benning, C.
Arch. Biochem. Biophys. 327, 27-34, 1996
A:Title: Functional Expression of Uridine 5'-diphospho-glucose 4-epimerase (EC 5.1.3.2)
A:Reference number: S62783; MUID:96201343; PMID:8615692
A:Accession: S62783
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-351 <DOE>
A:Cross-references: EMBL:Z54214; NID:g1143391; PIDN:CAA90941.1; PID:g1143392
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: galactose metabolism; isomerase

F:9-347/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 85.1%; Score 1571; DB 2; Length 351;
Best Local Similarity 85.6%; Pred. No. 3.3e-121;
Matches 297; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

QY 3 SSSOHILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 62
DB 4 SVEQNILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 63
QY 63 FTQGLNRDLEKLSKTTFDVAVHPAGLKAVASVAKPRRYFDNVLGTTINLYETMAK 122
DB 64 FNLGDLRNKGDIEKLFKQDFDAVHPAGLKAVGESVEKGRYFDNVLGTTINLYETMAK 123
QY 123 YNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKIL 182
DB 124 YNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKIL 183
QY 183 LRYFNPVGAHESKGLGDPKGI PNNLMPYIQQVAVGRLPELVYGHDPYTRDGSARDYI 242
DB 184 LRYFNPVGAHESKGLGDPKGI PNNLMPYIQQVAVGRLPELVYGHDPYTRDGSARDYI 243
QY 243 HVMDLADGHTAALRKLFTTENGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 302
DB 244 HVMDLADGHTAALRKLFTTENGCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCPRR 303
QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 349
DB 304 SGDATAVYASTERAELGKWKANYGVEMCRDQWNAKNNPWGYAGK 350

RESULT 4

D96657
hypothetical protein F16M1.9 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
A:Accession: D96657
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96657
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:Cross-references: GB:AB005173; NID:g10092244; PIDN:AAG12659.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
A:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 82.5%; Score 1522; DB 2; Length 353;
Best Local Similarity 81.4%; Pred. No. 3.5e-117;
Matches 285; Conservative 25; Mismatches 38; Indels 2; Gaps 1;

QY 3 SSSOHILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 62
DB 4 SVEQNILVTGAGFIGTHTVVQLLKAGFSVSIIDFNSVMEAVDRVQVGPLLSQNLQ 63
QY 63 FTQGLNRDLEKLSKTTFDVAVHPAGLKAVASVAKPRRYFDNVLGTTINLYETMAK 120
DB 64 FNLGDLRNKGDIEKLFKQDFDAVHPAGLKAVGESVGNPRYFDNVLGTTINLYETMAK 123
QY 121 AKYNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKI 180
DB 124 AKYNCKMVFSSATVYGOPEKIPCEBDFKLOAMPYGRTKLFLEIARDIQKAEPEWKI 183

QY 181 ILLRYFNPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRLTELVNHYGHDPYTRDGSALRD 240
 Db 184 ILLRYFNPVGAHSGRIGEDPKGIPNNLMPYIQQVAVGRLPELVNHYGHDPYTRDGSALRD 243
 QY 241 YHVMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 300
 Db 244 YHVMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 303
 QY 301 RRGDATEVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAGKP 350
 Db 304 RRGDATEVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAGKP 353

RESULT 5
 T10496
 UDPglucose 4-epimerase (EC 5.1.3.2) (clone GEPI42) - guar
 N;Alternate names: UDP-galactose 4-epimerase
 C;Species: *Cyamopsis tetragonoloba* (guar, cluster bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10496
 R;Brunstedt, J.; Joersbo, M.; Pedersen, S.G.; Marcussen, J.
 Plant Sci. 142, 147-154, 1999
 A;Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase ex
 A;Reference number: Z17058
 A;Accession: T10496
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-354 <BRU>
 A;Cross-references: EMBL:AJ005081
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;12-350/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 81.3%; Score 1500; DB 2; Length 354;
 Best Local Similarity 79.6%; Pred. No. 2.2e-115;
 Matches 277; Conservative 32; Mismatches 39; Indels 0; Gaps 0;

QY 2 VSSSHILVTGGAGFTGTHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNL 61
 Db 6 MASGETILVTGGAGFTGSHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNL 65
 QY 62 QFTQGLNRDRLLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNA 121
 Db 66 HFHGLDLNRHLDILFSKTKFDAVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNA 125
 QY 122 KYNCKKMFSSSATVYQPEKIPCEDEDFKLOANPYGRKLFLEETARDIOKAEPEWKII 181
 Db 126 KFNCKKLVISSSATVYQPEKIPCEDEDFKLOANPYGRKLFLEETARDIOKAEPEWKII 185
 QY 182 LLRYFNPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRLTELVNHYGHDPYTRDGSALRD 241
 Db 186 LLRYFNPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRLPELVNHYGHDPYTRDGSALRD 245
 QY 242 IHVMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 301
 Db 246 IHVMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCP 305
 QY 302 RRGDATEVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAGKP 349
 Db 306 RRGDATEVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAGKP 353

RESULT 6
 T10498
 UDPglucose 4-epimerase (EC 5.1.3.2) (clone GEPI48) - guar
 N;Alternate names: UDP-galactose 4-epimerase
 C;Species: *Cyamopsis tetragonoloba* (guar, cluster bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T10498
 R;Brunstedt, J.; Joersbo, M.; Pedersen, S.G.; Marcussen, J.
 Plant Sci. 142, 147-154, 1999
 A;Title: Isolation and expression of two cDNA clones encoding UDP-galactose epimerase ex

A;Reference number: Z17058
 A;Accession: T10498
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-350 <BRU>
 A;Cross-references: EMBL:AJ005082
 C;Function:
 A;Description: catalyzes the reversible epimerization of UDP-galactose and UDP-glucose
 A;Pathway: carbohydrate metabolism
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;6-342/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 67.5%; Score 1246; DB 2; Length 350;
 Best Local Similarity 67.5%; Pred. No. 1.6e-94;
 Matches 233; Conservative 44; Mismatches 66; Indels 2; Gaps 2;

QY 4 SSOHILVTGGAGFTGTHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQF 63
 Db 2 SSOHILVTGGAGFTGTHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQF 60
 QY 64 TQGLNRDRLLEKLFSTKTFDAVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNAKY 123
 Db 61 HKLDLRDRDALEKIFSSSTKTFDSVTHFAGLKAVASVAKPRYFDNLTGNTINLYEFNAKY 120
 QY 124 NCKKMFSSSATVYQPEKIPCEDEDFKLOANPYGRKLFLEETARDIOKAEPEWKIILL 183
 Db 121 GCKKMFSSSATVYQPEKIPCEDEDFKLOANPYGRKLFLEETARDIOKAEPEWKIILL 180
 QY 184 RYFNPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRLTELVNHYGHDPYTRDGSALRDYIH 243
 Db 181 RYFNPVGAHSGKLGEDPKGIPNNLMPYIQQVAVGRLPELVNHYGHDPYTRDGSALRDYIH 240
 QY 244 VMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCPRRP 303
 Db 241 VMDLADGHIALLKRLFTTENTIGCTAYNLGTRGTSVLEWTAPEKASGKKIPVKLCPRRP 299
 QY 304 GDATVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAG 348
 Db 300 GDATVYASTERAELGWKANYGVEMCRDQWNAKNNPWGAG 344

RESULT 7
 T10881
 UDPglucose 4-epimerase (EC 5.1.3.2) F8M12.10 - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jan-2000
 C;Accession: T01881
 R;Madsen, C.; Graves, T.; Cotton, M.; Modde, T.
 submitted to the EMBL Data Library, July 1998
 A;Description: The sequence of A. thaliana F8M12.
 A;Reference number: Z14450
 A;Accession: T01881
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-350 <MAD>
 A;Cross-references: EMBL:AF080118; NID:G3513725; PIDN:AAC33955.1; PID:G3513739
 A;Experimental source: cultivar Columbia
 C;Genetics:

A;Map position: 4
 A;Introns: 61/3; 78/2; 123/3; 157/3; 233/3; 260/1; 287/3; 318/2
 A;Note: F8M12.10
 C;Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
 C;Keywords: galactose metabolism; isomerase
 F;5-341/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 66.6%; Score 1229; DB 2; Length 350;
 Best Local Similarity 66.7%; Pred. No. 3.9e-93;
 Matches 228; Conservative 50; Mismatches 52; Indels 2; Gaps 2;

QY 5 SSOHILVTGGAGFTGTHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQFT 64
 Db 2 ARNVLSGGAGYIGSHVTVOLLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQFT 60

QY 65 QGDLNRDDLEKLFSTTTFDVAVIHAGLKAVAESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 61 QVDLRDRSALEKIFSTETKFDVAVIHAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 120
QY 125 CKKMVFSSSATVYGQPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIQKASPEWKIILLR 184
DB 121 CKNLVFGSSATVYGSPEKVPCTEPEPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 180
QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGVHDYPRDGSALRDYIHV 244
DB 181 YFNPGVAGHSPGDIAGEDPGIPNNLMPIYQOVAVGRPHLTVPFGNDYNTKDGTVGRDYIHV 240
QY 245 MDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 241 IDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 120
QY 305 DATEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 346
DB 300 DAEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 341

RESULT 8

T04291
probable UDPglucose 4-epimerase (EC 5.1.3.2) - Arabidopsis thaliana
N:Alternate names: protein F25124.170
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jan-2000
C:Accession: T04291
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15261
A:Accession: T04291
A:Molecule type: DNA
A:Residues: 1-351 <BEV>
A:Cross-references: EMBL:AL049525
A:Experimental source: cultivar Columbia; BAC clone F25124
C:Genetics:
A:Map position: 4
A:Introns: 62/3; 79/2; 124/3; 158/3; 234/3; 261/1; 288/3; 319/2
A:Note: F25124.170
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: isomerase
F:6-342/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 66.6%; Score 1229; DB 2; Length 351;
Best Local Similarity 66.7%; Pred. No. 3.9e-93; Mismatches 50; Indels 2; Gaps 2;
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;

QY 5 SQHILVTGGAGFTGTHVTVQVLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSQNLQFT 64
DB 3 ARNLVSGGAGYIGSTHTVLQLLGGYSVVVDNLNDSASVLSQVKKLAEE-HGERLSFH 61
QY 65 QGDLNRDDLEKLFSTTTFDVAVIHAGLKAVAESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 62 QVDLRDRSALEKIFSTETKFDVAVIHAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 121
QY 125 CKKMVFSSSATVYGQPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIQKASPEWKIILLR 184
DB 122 CKNLVFGSSATVYGSPEKVPCTEPEPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 181
QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGVHDYPRDGSALRDYIHV 244
DB 182 YFNPGVAGHSPGDIAGEDPGIPNNLMPIYQOVAVGRPHLTVPFGNDYNTKDGTVGRDYIHV 241
QY 245 MDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 242 IDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 120
QY 305 DATEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 346
DB 301 DAEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 342

RESULT 9

T08911
UDPglucose 4-epimerase (EC 5.1.3.2) T32A16.90 - Arabidopsis thaliana
N:Alternate names: protein T32A16.90
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000
C:Accession: T08911
R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Ma submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16518
A:Accession: T08911
A:Molecule type: DNA
A:Residues: 1-350 <BEV>
A:Cross-references: EMBL:AL078468; GSPDB:GN00062; ATSP:T32A16.90
A:Experimental source: cultivar Columbia; BAC clone T32A16
C:Genetics:
A:Gene: ATSP:T32A16.90
A:Map position: 4
A:Introns: 61/3; 78/2; 123/3; 157/3; 233/3; 260/1; 287/3; 318/2
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C:Keywords: galactose metabolism; isomerase
F:5-341/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 65.5%; Score 1208; DB 2; Length 350;

Best Local Similarity 65.5%; Pred. No. 2.1e-91; Mismatches 224; Conservative 49; Mismatches 67; Indels 2; Gaps 2;
QY 5 SQHILVTGGAGFTGTHVTVQVLLKAGFSVSIIDFNDSVMEAVDRVQVVGPLLSQNLQFT 64
DB 2 AKSVLVTGGAGYIGSTHTVLQLLGGYSVVVDNLNDSASVLSQVKKLAEE-HGERLSFH 60
QY 65 QGDLNRDDLEKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 61 QVDLRDRSALEKIFSTETKFDVAVIHAGLKAVGESVEKELLYNNLVGTITLLEVAQHG 120
QY 125 CKKMVFSSSATVYGQPEKIPCEBDFKLQAMNPGYGRTKLFLEIARDIQKASPEWKIILLR 184
DB 121 CKNLVFGSSATVYGSPEKVPCTEPEPISALNPGYGRTKLFIEICRDVYGSPEWKIILLR 180
QY 185 YFNPGVAGHSGKLGEDPKGIPNNLMPIYQOVAVGRLTSLNLYGVHDYPRDGSALRDYIHV 244
DB 181 YFNPGVAGHSPGDIAGEDPGIPNNLMPIYQOVAVGRPHLTVPFGNDYNTKDGTVGRDYIHV 240
QY 245 MDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 124
DB 241 MDLADGHIAALRKLFSTTTFDVAVIHAGLKAVESVAKPRRYFDENLVGTINLYEFMAKYN 120
QY 305 DATEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 346
DB 300 DAEVYASTERAELGKWKANYGVEMCRDWNWAKNPNWY 341

RESULT 10

D69628
UDPglucose 4-epimerase (EC 5.1.3.2) - Bacillus subtilis
N:Alternate names: UDPgalactose 4-epimerase
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: D69628
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Fritton, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Sero akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,


```
A;Gene: galE
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match      52.2%; Score 964; DB 2; Length 339;
Best Local Similarity 51.9%; Pred. No. 2.1e-71;
Matches 177; Conservative 59; Mismatches 99; Indels 6; Gaps 2;

Qy 6 QHILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNLQFTQ 65
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 EKILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNLQFTQ 58
Qy 66 GDLNRDLEKLFSTKTTFDVAVIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFMAKYN 125
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
59 ADIRDITDLRFKQEEPTGVHFAGLKAVGESTRIPLAYVDNNIAGTVSLLKAMEENN 118
Qy 126 KQWVFSSATVYQGPKEIPCEBDFKLOANVPYGRTKLFLEIARDIQKAPPEWKIILRY 185
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 KNIFSSATVYGDPTVPIEDFPLSVTNPYGRTKLMLEIITDIYKADSEWNVLLRY 178
Qy 186 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGSAIRDIYHM 245
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGSAIRDIYHM 238
Qy 246 DLADGHIAALRKLFTEENICTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 305
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 DLAKGHVAALKKI--QKSGNLVNYLGTGKYSVLEIIQNMEKAVGCRPIPIRIVERRPGD 296
Qy 306 ATEVYASTERAELGKWKANYGVEEMCRDWNWAKNPMGY 346
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 IAACYSDPKAKAELGWEAELDTQMCDAWRQSKHPNGF 337

RESULT 15
T19989
hypothetical protein C47B2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C;Accession: T19989
R;Karslaw, J.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z19208
A;Accession: T19989
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-347 <WLL>
A;Cross-references: EMBL:Z99709; PIDN: CAB16861.1; GSPDB: GN00019; CESP: C47B2.6
A;Experimental source: clone C47B2
C;Genetics:
A;Gene: CESP:C47B2.6
A;Map position: 1
A;Introns: 39/1; 115/3; 258/3; 296/3
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match      52.1%; Score 962; DB 2; Length 347;
Best Local Similarity 52.1%; Pred. No. 3.2e-71;
Matches 183; Conservative 59; Mismatches 94; Indels 16; Gaps 4;

Qy 7 HILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNL 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 HILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNL 57
Qy 61 LQFTQGLNRDLEKLFSTKTTFDVAVIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFM 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 VPFQVNDVCEAALEKVFSENKFDGIHLLAALKAVGESVAKPQYYSNNLVASLNTQMC 117
Qy 121 AKYNCKQWVFSSATVYQGPKEIPCEBDFKLOA----MNPYGRTKLFLEIARDIQKAP 176
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
118 LKYNKQWVFSSATVYQGPKEIPCEBDFKLOA----MNPYGRTKLFLEIARDIQKAP 175
Qy 177 EWKITLLRYNPNVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGS 236
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 EWNVLLRYNPNVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGS 235

A;Gene: galE
C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C;Keywords: isomerase

Query Match      52.4%; Score 966; DB 2; Length 339;
Best Local Similarity 51.9%; Pred. No. 1.5e-71;
Matches 177; Conservative 59; Mismatches 99; Indels 6; Gaps 2;

Qy 6 QHILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNLQFTQ 65
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 EKILVTGGAGFICTHTVTVQVLLKAGFSVSIIDNFDNSVMEAVDRVQVGVGPELLSQNLQFTQ 58
Qy 66 GDLNRDLEKLFSTKTTFDVAVIHFAGLKAVAESVAKPRYFDNVLVGTINLYEFMAKYN 125
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
59 ADIRDITDLRFKQEEPTGVHFAGLKAVGESTRIPLAYVDNNIAGTVSLLKAMEENN 118
Qy 126 KQWVFSSATVYQGPKEIPCEBDFKLOANVPYGRTKLFLEIARDIQKAPPEWKIILRY 185
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 KNIFSSATVYGDPTVPIEDFPLSVTNPYGRTKLMLEIITDIYKADSEWNVLLRY 178
Qy 186 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGSAIRDIYHM 245
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
179 FNPVGAHESKLGEDPKGIPNNLMPYIQVAVAGRLTELNVYGHDPYTRDGSAIRDIYHM 238
Qy 246 DLADGHIAALRKLFTEENICTAYNLGTRGTSVLEMTAFKASGKKIPVKLCPRRPGD 305
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
239 DLAKGHVAALKKI--QKSGNLVNYLGTGKYSVLEIIQNMEKAVGCRPIPIRIVERRPGD 296
Qy 306 ATEVYASTERAELGKWKANYGVEEMCRDWNWAKNPMGY 346
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
297 IAACYSDPKAKAELGWEAELDTQMCDAWRQSKHPNGF 337

RESULT 14
C98054
UDPglucose 4-epimerase (EC 5.1.3.2) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: C98054
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID: 21429245; PMID: 11544234
A;Accession: C98054
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-339 <KUR>
A;Cross-references: GB:AE007317; PIDN: AAL00264.1; PID: g15459117; GSPDB: GN00174
C;Genetics:
```

QY	237	AIRDYIHVMDLADCHIAALRKLFTTENIGCTAYNLGTGRGTSVLEWVAFEXASGKKIPV	296
Db	236	GVEDYIHVVDLAKGHVKAFDRIKTVGNIGTEIYNLGTGVGYSVRQWVDALKKVSGRDIPV	295
QY	297	KLCPRRPGDATEVYASTERAEKELGHWKANYGVEEMCRDOWNWAKNPNQYA	347
Db	296	KIGVPRPGDVASYCDPSLAQEXLWRAETGLEEMCADLWNQTKNPOGFS	346

Search completed: March 22, 2004, 21:56:00
Job time : 28 secs

Db 1 MVASSCKILVTGAGFIGHTVTVQLLNGFNVSIIIDNFDSNMEAVREVVWGSNLSON 60
 QY 61 LQFTQGLDRNRDLEKLSKTTTDAVIFHAGLKAVASVAKPRYFDFNLVGTINLYEFM 120
 Db 61 LEFTLGLDRNRDLEKLSKSKFDVIFHAGLKAVGESVENPRYFDFNLVGTINLYEFM 120
 QY 121 AKYNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIQKAEPEWKI 180
 Db 121 AKHNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIQKAEPEWKI 180
 QY 181 ILLRYFNPVGAHESKLGEDPKGIPNNIMPYIQQVAVGRLTAVVGHYDTPRDSAIRD 240
 Db 181 VLLRYFNPVGAHESKLGEDPKGIPNNIMPYIQQVAVGRLTAVVGHYDTPRDSAIRD 240
 QY 241 YHVNMDLADGHIAALRKLFTEENICTAYNLTGRTGTSVLEMTVAFKASGKKIPVKLCP 300
 Db 241 YHVNMDLADGHIAALRKLFTEENICTAYNLTGRTGTSVLEMTVAFKASGKKIALKLC 300
 QY 301 RPPGDATEVYASTAEKELGWKANYGVEMCRDQWKNWKNPNWYAGKP 350
 Db 301 RPPGDATEVYASTAEKELGWKAKYGVEMCRDQWKNWKNPNWYSGKP 350

RESULT 2

GAEL ARATH STANDARD; PRT; 351 AA.
 AC Q42605, Q9LPX1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 GN AT1G12780 OR F13K23.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI TaxID=3702;
 RN [1]
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=cv Columbia;
 RX MEDLINE=96201343; PubMed=8615692;
 RA Doermann P., Benning C.;
 RT "Functional expression of uridine 5'-diphospho-glucose 4-epimerase
 RT (EC 5.1.3.2) from Arabidopsis thaliana in Saccharomyces cerevisiae and
 RT Escherichia coli.";
 RL Arch. Biochem. Biophys. 327:27-34 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Atayjo R., Bowman C.L., Brooks S.V.,
 RA Buehler E., Chan A., Chao O., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana".
 RT Nature 408:816-920 (2000).
 CC -|- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -|- COFACTOR: Does not seem to necessitate NAD for activity.
 CC -|- PATHWAY: Galactose metabolism; third step.

CC -|- PATHWAY: Biosynthesis of UDP-galactose as precursor of
 CC galactolipids and cell wall polysaccharides.
 CC -|- TISSUE SPECIFICITY: Ubiquitously expressed. Highest expression
 CC in stems and roots.
 CC -|- MISCELLANEOUS: Shows a broad pH optimum of 7.0 to 9.5 and a Km of
 CC 0.11 mM.
 CC -|- SIMILARITY: Belongs to the sugar epimerase family.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL; Z54214; CAA90941.1; -
 CC EMBL; AC012187; AAF78483.1; -
 CC PIR; B6261; B6261.
 CC PIR; S62783; S62783.
 CC HSP; P09147; IXEL.
 CC InterPro; IPR001509; Epimerase_Dh.
 CC InterPro; IPR005886; GalE.
 CC InterPro; IPR000205; NAD BS.
 CC InterPro; IPR008089; Nuc_sugar_epim.
 CC Pfam; PF01370; Epimerase_1.
 CC PRINTS; PR01713; NUCBPIMERASE.
 CC TIGRfams; TIGR01179; GalE; 1.
 KW isomerase; NAD; Galactose metabolism; Multigene family.
 FT NP BIND 8 39 NAD (POTENTIAL).
 FT CONFLICT 102 103 NP -> KG (IN REF. 1).
 FT CONFLICT 341 341 N -> F (IN REF. 1).
 SQ SEQUENCE 351 AA; 39157 MW; FDB640FB26A1DD1D CRC64;
 Query Match 85.5%; Score 1578; DB 1; Length 351;
 Best Local Similarity 85.6%; Pred. No. 3e-118;
 Matches 237; Conservative 16; Mismatches 34; Indels 0; Gaps 0;
 QY 3 SSSQHLVTGGAGFIGHTVTVQLLKGAFSVSIIDNFDSNMEAVREVVWGSNLSONIQ 62
 Db 4 SVEQNLTVTGGAGFIGHTVTVQLLKDGFKVSIIDNFDSNMEAVREVVWGSNLSONIQ 63
 QY 63 FTQGLDRNRDLEKLSKTTTDAVIFHAGLKAVASVAKPRYFDFNLVGTINLYEFMAK 122
 Db 64 FNLGLDRNRDLEKLSKQKQFQDAVIFHAGLKAVGESVENPRYFDFNLVGTINLYETMAK 123
 QY 123 YNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIQKAEPEWKIL 182
 Db 124 YNCKMVFSSSATVYQGPKEIPCEBDFKLQANMPYGRTKLFLEETARDIQKAEPEWKIL 183
 QY 183 LRYFNPVGAHESKLGEDPKGIPNNIMPYIQQVAVGRLTAVVGHYDTPRDSAIRDYI 242
 Db 184 LRYFNPVGAHESKLGEDPKGIPNNIMPYIQQVAVGRLTAVVGHYDTPRDSAIRDYI 243
 QY 243 HVMDLADGHIAALRKLFTEENICTAYNLTGRTGTSVLEMTVAFKASGKKIPVKLCPR 302
 Db 244 HVMDLADGHIAALRKLFADPKTGCTAYNLTGRTGTSVLEMTVAFKASGKKIPVKLCPR 303
 QY 303 PGDATEVYASTAEKELGWKANYGVEMCRDQWKNWKNPNWYAGK 349
 Db 304 SGDATEVYASTAEKELGWKANYGVEMCRDQWKNWKNPNWYQNK 350
 RESULT 3
 GAEL CYATE STANDARD; PRT; 354 AA.
 AC O65780;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase GEPI42 (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 OS Cyamopsis tetragonoloba (Guar) (Cluster bean).

Qy	248	ADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLVEMTWTAPEKASGKKI.PVKLCPPRPGDAT	307
Db	239	AEHGVTAKLEKVL--NSTGADAYNLGTGTSVLVEMTWKAFKVKSVKGYFYFADRRPGDIA	296
Qy	308	EVVASTERAKELGKWKANYGVVEEMCRDQWNAKNPMGY	346
Db	297	TCFADPAKAKRELGWKRELGLEWCADSWNQSSNVNGY	335

RESULT	ID	GALE_MOUSE	STANDARD;	PRT;	347	AA.
AC	Q8R059;					
DT	10-OCT-2003	(Rel. 42, Created)				
DT	10-OCT-2003	(Rel. 42, Last sequence update)				
DT	10-OCT-2003	(Rel. 42, Last annotation update)				
DE	UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-					
DE	galactose 4-epimerase).					
GN	GALE.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
RN	[1]					
RC	SEQUENCE FROM N.A.					
RC	TISSUE=Colon;					
RC	MEDLINE=2238257; PubMed=12477932;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,					
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,					
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,					
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,					
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,					
RA	Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,					
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,					
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,					
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,					
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,					
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,					
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,					
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,					
RA	Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;					
RT	"Generation and initial analysis of more than 15,000 full-length					
RT	human and mouse cDNA sequences."					
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).					
CC	-1- FUNCTION: Catalyzes two distinct but analogous reactions: the					
CC	epimerization of UDP-glucose to UDP-galactose and the					
CC	epimerization of UDP-N-acetylglucosamine to UDP-N-					
CC	acetylglactosamine.					
CC	-1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.					
CC	-1- COFACTOR: NAD.					
CC	-1- PATHWAY: Galactose metabolism; third step.					
CC	-1- SIMILARITY: Belongs to the sugar epimerase family.					
CC						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/).					
CC	or send an email to license@ebi-sib.ch .					
CC						
CC	ENBL; BC027438; AAR27438.1; -					
DR	MGD; MGI:1921496; Gale.					
DR	InterPro; IPR001509; Epimerase_Dh.					
DR	InterPro; IPR005886; Gale.					
DR	InterPro; IPR008089; Nuc_sugar_epim.					
DR	Pfam; PF01370; Epimerase; 1.					
DR	PRINTS; PR01713; NUCEPIMERASE.					
DR	TIGRFAMs; TIGR01179; gale; 1.					
KW	Isomerase; NAD; Galactose metabolism.					

QY 120 MAKYNCKMVFSSATVYGOPEKIPCEEDFKLQAMNYPGRTKFLBEIARDIOKASPEWK 179
 Db 119 MRANGVSLVFSSATVYGPVPSASGRPPHRCGTPYKSKFFIEMIQDLCRADTAWN 178
 QY 180 IILLRYNPVGAHESKGLGDPKGIPINNLMPYIQQVAVGRLTELNVYGHDPYPRDGSAIR 239
 Db 179 AVLLRYPIPIGAHRSARIGEDPGIPINNLMPYISQVAGREALNVFGDDYATEDGTGYR 238
 QY 240 DYIHMVLAUGHIAALRLFTTNIENIGCTANLGTGRTSVLEMTAFKASGKKIPVKJC 299
 Db 239 DYIHMVLAUGHIAALRLFTTNIENIGCTANLGTGRTSVLEMTAFKASGKKIPVKJC 299
 QY 300 PRPFGDATEVYASTERAELKGMKANYGVEMCRDQNNWAKNPNWG 345
 Db 297 ARREGDVAACYPNSLAHEELGWTALGLDRNCEDLWRWQKNPSG 342

RESULT 11
 ID_GALE_NEIMA STANDARD; PRT; 339 AA.
 AC P56957;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 galactose 4-epimerase).
 GN GALE OR NMA0203.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
 GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
 WHICH ARE IMPORTANT FOR PATHOGENESIS.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar epimerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL162752; CAB83517.1; -.
 DR F82014; F82014.
 DR HSPF; P09147; IXEL.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GALE.
 DR Pfam; PF01370; Epimerase; 1.
 DR TIGRFAMs; TIGR01179; gale; 1.
 KW Isomerase; NAD; Galactose metabolism; Complete proteome.
 FT NP_BIND 339 34
 SQ SEQUENCE 339 AA; 37066 MW; A68DB884417B0C1C CRC64;

Query Match 50.3%; Score 928.5; DB 1; Length 339;

Best Local Similarity 52.5%; Pred. No. 1.4e-66;
 Matches 180; Conservative 58; Mismatches 96; Indels 9; Gaps 4;

QY 6 OHILVTGAGTGTHTVVOLLKAGFSVSIIDFNFSNVMEAVDRVQVVGPLLSQNLQFTQ 65
 Db 2 KNILVTGAGTGTHTVVOLLKAGFSVSIIDFNFSNVMEAVDRVQVVGPLLSQNLQFTQ 65
 QY 66 GLNRDRLKLFKSTTDFAVIHFAGLKAFAESVAKPRYDFNLTGNTINLYEFNAKNC 125
 Db 58 GDIRDREILRIFAENRIDSIVHFAGLKAFAESVAKPRYDFNLTGNTINLYEFNAKNC 117
 QY 126 KMVPSSSATVYGOPEKIPCEEDFKL-QAMNYPGRTKFLBEIARDIOKASPEWKIILLR 184
 Db 118 PSIVPSSSATVYGOPEKIPCEEDFKL-QAMNYPGRTKFLBEIARDIOKASPEWKIILLR 177
 QY 185 YFNPVGAHESKGLGDPKGIPINNLMPYIQQVAVGRLTELNVYGHDPYPRDGSAIRDYIHV 244
 Db 178 YFNPVGAHESKGLGDPKGIPINNLMPYIQQVAVGRLTELNVYGHDPYPRDGSAIRDYIHV 237
 QY 245 MDLADGHIAALRLKLTFTTNI-GCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCRPP 303
 Db 238 MDLADGHIAALRLKLTFTTNI-GCTAYNLGTGRGTSVLEMTAFKASGKKIPVKLCRPP 294
 QY 304 GDATEVYASTERAELKGMKANYGVEMCRDQNNWAKNPNWG 346
 Db 295 GDLACFYADPSYTKAQIGWQTDRLAQWEDSWRWSNNPNGY 337

RESULT 12
 ID_GALE_NEIMC STANDARD; PRT; 339 AA.
 AC P56986; Q59617; Q59624;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 galactose 4-epimerase).
 GN GALE.
 OS Neisseria meningitidis (serogroup C).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=135720;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FAM20 / Serogroup C;
 RX MEDLINE=95310004; PubMed=7790063;
 RA Lee F.K., Stephens D.S., Gibson B.W., Engstrom J.J., Zhou D.,
 RA Apicella M.A.;
 RT "Microheterogeneity of Neisseria lipooligosaccharide: analysis of a
 UDP-glucose 4-epimerase mutant of Neisseria meningitidis NMB.";
 RL Infect. Immun. 63:2508-2515(1995).
 CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
 GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
 WHICH ARE IMPORTANT FOR PATHOGENESIS.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar epimerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U19895; AAA86716.1; -.
 DR HSPF; P09147; IXEL.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GALE.
 DR Pfam; PF01370; Epimerase; 1.

DR TIGRFAMS; TIGR01179; gale; 1.
 KW isomerase; NAD; Galactose metabolism.
 FT NP BIND 3 34 NAD (POTENTIAL).
 SQ SEQUENCE 339 AA; 37030 MW; 7DC4E4ACBE046397 CRC64;

Query Match
 Best Local Similarity 50.3%; Score 928.5; DB 1; Length 339;
 Matches 180; Conservative 57; Mismatches 95; Indels 9; Gaps 4;

QY 8 ILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGPLLSQNLQFTQGD 67
 DB 4 ILVTGGFTGFTSVLSLKGSHQVVDLNLCSNINILPRLKTIIG-----QEIPFYQGD 59
 QY 68 LNRDDLEKLFSTTDFAVIHFAFLKAVASVAKPRYDFNLTNGTINLYEFMAKYNCKK 127
 DB 60 IRDRLELRIFAENRDSVHFAFLKAVGESVABPMKYDNNVSGSLVAEEMARAGVFK 119
 QY 128 MVFSSSATVYGOPEKIPCEBDFKL-QAMNPYGRTKLFLLEIARDIOKAEPEWKILLRYF 186
 DB 120 IVFSSSATVYGDGKVPYTEDMQPGDTTSPYGTSKSWVERILSDIOKADPRMSVILLRYF 179
 QY 187 NPVGAHESKGLGSDPKGIPNNLMPYIQQVAVGRITELNLYCHDYPTTRDGSARDYIHVMD 246
 DB 180 NPVGAHESGLIGEDPENGIPNNLLPYISQVAGKLAQLSVFGSDYDTHDGTGVRDYIHVMD 239
 QY 247 LADGHIAALRKLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRPGD 305
 DB 240 LAEGHVAAMQ--AKSNVAGTHLNLGSRASSVLEIRAFEAASGLTIPIYEVKPRRAGD 296
 QY 306 ATEVYASTERAEKELGKANYGVEMCRDOWNAKNNPWGY 346
 DB 297 LACFYADPSYAKAIGQWQTQDILQTMEDSWRVVSNPNGY 337

RESULT 13
 GALE_HAEIN STANDARD; PRT; 338 AA.
 AC P24325;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 GN GALE OR H10351.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RM 7004 / Serotype B;
 RX MEDLINE=92065797; PubMed=7542800;
 RA Maskell D.J., Szabo M.J., Butler P.D., Williams A.E., Moxon E.R.;
 RT "Molecular analysis of a complex locus from Haemophilus influenzae
 RT involved in phase-variable lipopolysaccharide biosynthesis.";
 RL Mol. Microbiol. 5:1013-1022 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Karlvage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512 (1995).
 CC -1- FUNCTION: BY CONTROLLING THE INTERNAL GALACTOSE CONCENTRATION,

CC IT MAY BE LINKED TO THE BIOSYNTHESIS OF LIPOPOLYSACCHARIDE
 CC SURFACE MOLECULES, WHICH ARE IMPORTANT FOR THE PATHOGENESIS OF
 CC H. INFLUENZAE.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COFACTOR: NAD
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the sugar epimerase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X57315; CAA40568.1; -;
 CC EMBL; U32719; AAC22012.1; -;
 CC PIR; A64063; A64063.
 CC HSP; P09147; IXL.
 CC TIGR; H10351; -;
 CC InterPro; IPR001509; Epimerase_Dh.
 CC InterPro; IPR005886; Gale.
 CC InterPro; IPR000205; NAD_BS.
 CC InterPro; IPR008089; Nuc_sugar_epim.
 CC Pfam; PF01370; Epimerase; 1.
 CC PRINTS; PR01713; NUCEPIMERASE.
 CC TIGRFAMS; TIGR01179; gale; 1.
 KW isomerase; NAD; Galactose metabolism; Complete proteome.
 FT NP BIND 2 33 NAD (POTENTIAL).
 FT VARIANT 285 285 A -> P (IN STRAIN RM 7004).
 FT VARIANT 333 333 P -> S (IN STRAIN RM 7004).
 SQ SEQUENCE 338 AA; 37165 MW; 449934B0E21C8A56 CRC64;

Query Match 49.9%; Score 920.5; DB 1; Length 338;
 Best Local Similarity 53.2%; Pred. No. 6.2e-66;
 Matches 181; Conservative 56; Mismatches 96; Indels 7; Gaps 3;

QY 8 ILVTGGAGFTGTHVQLLKAGFSVSIIDNFNSVMEAVDRVQVQVGPLLSQNLQFTQGD 67
 DB 3 ILVTGGAGYIGSHVIVVELLVGKVVVDNLCSNIPKSLRVKQITG----KEAKFEGD 58
 QY 68 LNRDDLEKLFSTTDFAVIHFAFLKAVASVAKPRYDFNLTNGTINLYEFMAKYNCKK 127
 DB 59 ILDRALLQKIPAEINSEINSHFAFLKAVGESVQKPYEYMNNAVAGTLVLQEMKKAGVWN 118
 QY 128 MVFSSSATVYGOPEKIPCEBDFKL-QAMNPYGRTKLFLLEIARDIOKAEPEWKILLRYF 186
 DB 119 FVFSSSATVYGDPKIIPITDECEVGGTNPYGTSKYVVEQLRDTAKAEKFSNTILRYF 178
 QY 187 NPVGAHESKGLGSDPKGIPNNLMPYIQQVAVGRITELNLYCHDYPTTRDGSARDYIHVMD 246
 DB 179 NPVGAHESGLIGEDPENGIPNNLLPYISQVAGKLAQLSVFGSDYDTHDGTGVRDYIHVMD 238
 QY 247 LADGHIAALRKLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRPGDA 306
 DB 239 LAVGHLXALQR--HENDAGLHLYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRPGDA 296
 QY 307 TEVYASTERAEKELGKANYGVEMCRDOWNAKNNPWGY 346
 DB 297 ATCISDPSLAKELGKVAERGLKEMQDWTNNWQKNPKGY 336

RESULT 14
 GALE_DROME STANDARD; PRT; 350 AA.
 AC Q9W0E5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).

CG12030.
 GN Drosophila melanogaster (fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "the genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan X.H.,
 RA Rubin G.M., Celniker S.E.,
 RT "A Drosophila full-length cDNA resource."
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -1- FUNCTION: Catalyzes two distinct but analogous reactions: the
 CC epimerization of UDP-glucose to UDP-galactose and the
 CC epimerization of UDP-N-acetylglucosamine to UDP-N-
 CC acetylglucosamine (by similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -1- COFACTOR: NAD (by similarity).
 CC -1- PATHWAY: Galactose metabolism; third step.
 CC -1- SIMILARITY: Belongs to the sugar epimerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE003469; AAF47398.1; -
 DR EMBL; AY058582; AAL1381.1; -
 DR HSS; F09147; 1XEL.
 DR FlyBase; FBgn0035147; CG12030.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; Gale.
 DR InterPro; IPR008089; Nuc_sugar_epim.
 DR Pfam; PF01370; Epimerase; 1.
 DR PRINTS; PR01713; NUCEPIMERASE.
 DR TIGRfams; TIGR01179; gale; 1.
 KW Isomerase; NAD; Galactose metabolism.
 FT NP_BIND 5 36 NAD (POTENTIAL).
 SQ SEQUENCE 350 AA; 38697 MW; 0F14EDCC089ADFS CRC64;
 Query Match 49.8%; Score 918.5; DB 1; Length 350;
 Best Local Similarity 50.7%; Pred. No. 9.3e-66;
 Matches 177; Conservative 57; Mismatches 104; Indels 11; Gaps 3;
 QY 8 ILVTGGAGFITHVTVQLLKAGFSVSIIDNFDNS-----VMEAVDRVRQVVGPIISQNL 61
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 61
 6 VLVTGGAGYIGSHTVLEMLNAGYVVICVDNLICNAYSSGAKLPEALSRYQIEITG---KKV 61
 QY 62 OPTQGLDNRDLLEKLFSTKTFDAVHPAGLKVAESVAKPRRYFDENLVGTINLYEFMA 121
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 121
 62 NFYRVDIIDRQVRSVFOEHKIDVAFALKAVGESCRIPLOYHNNMTGTNVLLEMA 121
 QY 122 KYNCKMVFSSSATYGOEPIPCBEDPKL-QANNPYGRTKLFLEEIARDIOKAEPEWKI 180
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
 122 DNNVFKFYSSSATYVGPKEPLVTESHPTGNCCTSPYKTKYFTTEILKDLCKSDKRWAV 181
 QY 181 ILLRYFENVGHAESKGLGDPKIPNNIMPYIQOVAVGRLTENVYGHDPTRDGSALRD 240
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
 182 VSLRYFENVGHAISGRIGEDPENGPNLMPYIAQVAVGRRPSLSVYSGDFTHTDGTGVRD 241
 QY 241 YHVMDLADGHIAALRKLFITTENIGCTAYNIGTGRGTSVLEMTAFKASGKKIPVKLC 300
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
 242 YIHVLDLAEGHVKALDKLRNIAETGFFAYNLGTGVYSLVDWKAFKASGKKVNYTLVD 301
 QY 301 RPPGDATVYASTAEAEKELCKANYGVEMRCRDQWVWAKNPNWYAGK 349
 DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 349
 302 RRSQGVATCYADATLADKKLGWKAERGIDKMCEDTWRQSQNPNGYANK 350
 RESULT 15
 GALE_NEIMB
 ID GALE_NEIMB STANDARD; PRT; 339 AA.
 AC P56365; Q59617; Q59624;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
 DE galactose 4-epimerase).
 DE GALE OR NMB0064.
 GN Neisseria meningitidis (serogroup B).
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=95020536; PubMed=7934827;
 RA Jennings M.P., van der Ley P., Wilks K.E., Maskell D.J., Poolman J.T.,
 RA Moxon E.R.,
 RT "Cloning and molecular analysis of the gale gene of Neisseria
 RT meningitidis and its role in lipopolysaccharide biosynthesis."
 RL Mol. Microbiol. 10:361-369(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scariato V., Maignani V., Fizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.,
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN THE INCORPORATION OF
CC GALACTOSE INTO MENINGOCOCCAL LIPOPOLYSACCHARIDE SURFACE MOLECULES,
CC WHICH ARE IMPORTANT FOR PATHOGENESIS.
CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: Galactose metabolism; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the sugar epimerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20495; AAA6535.1; -.
DR EMBL; AE002366; AAF40532.1; -.
DR PIR; S39638; S39638.
DR HSP; P09147; 1XEL.
DR TIGR; NMB0064; -.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GalE.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; GalE; 1.
KW Isomerase; NAD; Galactose metabolism; Complete proteome.
FT NP BIND 3 34 NAD (POTENTIAL).
SQ SEQUENCE 339 AA; 37062 MW; 7D702C44F07DA99B CRC64;

Query Match 49.7%; Score 916.5; DB 1; Length 339;
Best Local Similarity 52.2%; Pred. No. 1.3e-65;
Matches 178; Conservative 56; Mismatches 98; Indels 9; Gaps 4;

QY 8 ILVTGAGFIGHTVVLKAGFSVSIIDNPNVMEAVDRVQVVGPELLSQNLOFTQGD 67
DB 4 ILVTGAGFIGHTVVLKAGFSVSIIDNPNVMEAVDRVQVVGPELLSQNLOFTQGD 59
QY 68 LNRDLEKLFKFTFDVAVIHAGLKAIVSAKPRPRYDFNLVGTINLYEFMAKYNCK 127
DB 60 IRDREILRFIAENRIDSVIHAGLKAIVSAKPRPRYDFNLVGTINLYEFMAKYNCK 119
QY 128 MVFSSSATVYGQEKIPCEEDFKL-QAMNPYGRTKLFLFEIARDIQKAEPEWKILLRYF 186
DB 120 IVFSSSATVYGQEKIPCEEDFKL-QAMNPYGRTKLFLFEIARDIQKAEPEWKILLRYF 179
QY 187 NPVGAHESKGLGDPKGIQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 246
DB 180 NPVGAHESKGLGDPKGIQVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 239
QY 247 LADGHTAALKLFTTNI-GCTAYNLGTGRTSVLEMTAFKASKKIPVKLCRRPFGD 305
DB 240 LAEGHVAWQ---AKSNVAGTHLLNGSRASVLEIIFAFASGLTIPYEVKPRAGD 296
QY 306 ATEVYASTERAEKELGWKANYGVEMCRDQWNAKNPNWGY 346
DB 297 LACFYADPSYTKAIGWQTDRLTQWEDSWRWVNNPNWGY 337

Search completed: March 22, 2004, 21:54:12
Job time : 20 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 21:40:15 ; Search time 67 Seconds
(without alignments)
1648.229 Million cell updates/sec

Title: US-09-913-064A-14

Perfect score: 1845

Sequence: 1 MVSSSHQHLVLTGAGFPGTH.....MCRDQNWAKNPNWYAGKP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1530	82.9	351	10 Q8LDN8	Q8LDN8 arabidopsis
2	1522	82.5	353	10 Q9CAM5	Q9CAM5 arabidopsis
3	1328	72.0	360	10 Q8H0B7	Q8H0B7 oryza sativ
4	1229	66.6	348	10 Q8LEA9	Q8LEA9 arabidopsis
5	1229	66.6	351	10 Q8VZ26	Q8VZ26 arabidopsis
6	1222	66.2	348	10 Q9C7W7	Q9C7W7 arabidopsis
7	1203	65.2	350	10 Q8LFW1	Q8LFW1 arabidopsis
8	1196.5	64.9	355	10 Q7XZQ2	Q7XZQ2 zea mays (m
9	1180	64.0	354	10 Q8LNZ3	Q8LNZ3 oryza sativ
10	1159.5	62.8	408	10 Q8H931	Q8H931 oryza sativ
11	1157.5	62.7	447	10 Q9SGX0	Q9SGX0 arabidopsis
12	1046	56.7	338	10 Q81414	Q81414 bacillus ce
13	1023	55.4	338	16 Q81JK4	Q81JK4 bacillus an
14	999.5	54.2	348	13 Q7Z2M6	Q7Z2M6 xenopus lae
15	998	54.1	337	16 Q8SGE0	Q8SGE0 shewanella
16	983	53.3	338	16 Q9KLH0	Q9KLH0 vibrio chol

17	980.5	53.1	339	16 Q87M58	Q87M58 vibrio para
18	977.5	53.0	335	16 Q8DBP1	Q8DBP1 vibrio vuln
19	973	52.7	339	2 Q83W11	Q83W11 streptococc
20	972	52.7	336	16 Q8DCQ4	Q8DCQ4 vibrio vuln
21	971.5	52.7	335	16 Q8VH83	Q8VH83 brucella me
22	970.5	52.6	335	2 Q54385	Q54385 brucella ab
23	970.5	52.6	335	16 Q8GDM3	Q8GDM3 brucella su
24	968.5	52.5	344	5 Q86153	Q86153 dictyosteli
25	966	52.4	339	16 Q97PK2	Q97PK2 streptococc
26	964	52.2	339	16 Q8DNY6	Q8DNY6 streptococc
27	962	52.1	347	5 Q62107	Q62107 caenorhabdi
28	961	52.1	338	16 Q87HS7	Q87HS7 vibrio para
29	951	51.5	364	2 Q5F8B2	Q5F8B2 moraxella c
30	950.5	51.5	335	16 Q82WJ5	Q82WJ5 nitrosomon
31	944	51.2	340	2 Q69132	Q69132 burkholderi
32	943	51.1	364	2 Q9KHV0	Q9KHV0 moraxella c
33	941	51.0	336	16 Q97P16	Q97P16 streptococc
34	927.5	50.3	339	2 Q9X3S6	Q9X3S6 neisseria m
35	927	50.2	340	16 Q8G3V9	Q8G3V9 bifidobacte
36	921.5	49.9	348	16 Q7V4P1	Q7V4P1 prochloroco
37	916.5	49.7	339	2 Q51148	Q51148 neisseria m
38	914	49.5	338	16 Q8XN29	Q8XN29 clostridium
39	913.5	49.5	338	2 Q937X4	Q937X4 edwardsiell
40	911	49.4	341	16 Q8XQW7	Q8XQW7 ralstonia s
41	908	49.2	326	16 Q8DNL9	Q8DNL9 streptococc
42	894.5	48.5	336	2 Q60109	Q60109 versinia en
43	894.5	48.5	338	2 P95527	P95527 pasteurella
44	894.5	48.5	338	16 Q8X942	Q8X942 escherichia
45	894.5	48.5	338	16 Q8FJR9	Q8FJR9 escherichia

ALIGNMENTS

RESULT 1

ID	Q8LDN8	PRELIMINARY;	PRT;	351 AA.
AC	Q8LDN8;			
DT	01-OCT-2002 (TRENBLrel. 22, Created)			
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)			
DE	Uridine diphosphate glucose epimerase, putative.			
GN	AtUG63180/F16M19.8.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosidII; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_taxonomy:3702;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,			
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;			
RT	"Full-length messenger RNA sequences greatly improve genome			
RT	annotation.";			
RL	Genome Biol. 0:0-0(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,			
RA	Feldmann K.;			
RT	"Full-length cDNA from Arabidopsis thaliana.";			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,			
RA	Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,			
RA	Havashizaki Y., Shinozaki K.;			
RT	"Arabidopsis thaliana full-length cDNA.";			
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBSJ databases.			
DR	EMBL; AY085887; AAMG3099.1; -			
DR	EMBL; AK117913; BAC42551.1; -			
DR	GO; GO:0003824; P: catalytic activity; IEA.			
DR	GO; GO:0003978; F: UDP-glucose 4-epimerase activity; IEA.			

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.
 DR InterPro; IPR000205; NAD_BS.
 DR InterPro; IPR008089; Nuc_sugar_epim.
 DR Pfam; PF01370; Epimerase; 1.
 DR PRINTS; PR01713; NUCEPIMERASE.
 DR TIGRFAMs; TIGR01179; GalE; 1.
 SQ SEQUENCE 351 AA; 38910 MW; 314BC1FABDD091EF CRC64;

Query Match 82.9%; Score 1530; DB 10; Length 351;
 Best Local Similarity 81.9%; Pred. No. 5.3e-111;
 Matches 285; Conservative 24; Mismatches 39; Indels 0; Gaps 0;

QY 3 SSSQHILVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSONLQ 62
 Db 4 SVEQNLVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSONLQ 63

QY 63 FTQGLNRDDLEKLFSTKTPDAVTHFAGLKAVAESVAKPRYPDFNVLVGTINLYEFMAK 122
 Db 64 FNLGLNRKGDIEKLFSTKTPDAVTHFAGLKAVGESVGNPRYPDFNVLVGTINLYEFMAK 123

QY 123 YNCKMWFSSATVYGOPEKIPCEDEKLOANMPYGRKLFLEIARDIQKAEPEWKIL 182
 Db 124 YNCKMWFSSATVYGOPEKIPCEDEKLOANMPYGRKLFLEIARDIQKAEPEWKIL 183

QY 183 LAYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRDYI 242
 Db 184 LAYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLPELVNFGHDPYTRDGSAIRDYI 243

QY 243 HWMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCPRR 302
 Db 244 HWMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLCPRR 303

QY 303 PGDATEVYASTERAELGKWKANYGVEMCRDQWNAKNPWGAGKP 350
 Db 304 AGDATAVYASTQKAEKELGKWKANYGVEMCRDQWNAKNPWGAGKP 351

RESULT 2
 Q9CAM5 PRELIMINARY; PRT; 353 AA.
 AC Q9CAM5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Uridine diphosphate glucose epimerase, putative, 80611-78786.
 GN F16M19.8.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.F.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koc H.L., Kreneskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maity B., Marziani A.,
 RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Taiton L.J., Tambunga G., Toriumi M.J., Town C.D.,

RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 EL Nature 408:816-820 (2000).
 DR EMBL; AC010795; AAG51599.1; -.
 DR PIR; D96657; D96657.
 DR HSP; P09147; IXEL.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.
 DR InterPro; IPR000205; NAD_BS.
 DR InterPro; IPR008089; Nuc_sugar_epim.
 DR Pfam; PF01370; Epimerase; 1.
 DR PRINTS; PR01713; NUCEPIMERASE.
 DR TIGRFAMs; TIGR01179; GalE; 1.
 SQ SEQUENCE 353 AA; 39180 MW; 398900C063C2C337 CRC64;

Query Match 82.5%; Score 1522; DB 10; Length 353;
 Best Local Similarity 81.4%; Pred. No. 2.2e-110;
 Matches 285; Conservative 25; Mismatches 38; Indels 2; Gaps 1;

QY 3 SSSQHILVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSONLQ 62
 Db 4 SVEQNLVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDSVMEAVDRVQVVGPLLSONLQ 63

QY 63 FTQGLNRDDLEKLFSS--KTFDPAVTHFAGLKAVAESVAKPRYPDFNVLVGTINLYEFM 120
 Db 64 FNLGLNRKGDIEKLFSTKTPDAVTHFAGLKAVGESVGNPRYPDFNVLVGTINLYEFM 123

QY 121 AKYCKMWFSSATVYGOPEKIPCEDEKLOANMPYGRKLFLEIARDIQKAEPEWKI 180
 Db 124 AKYCKMWFSSATVYGOPEKIPCEDEKLOANMPYGRKLFLEIARDIQKAEPEWKI 183

QY 181 ILLRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLTELNVYGHDPYTRDGSAIRD 240
 Db 184 ILLRYFNPVGAHESGKLGEDPKGIPNNLMPYIQQAVAGRLPELVNFGHDPYTRDGSAIRD 243

QY 241 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLC 300
 Db 244 YHVMDLADGHIAALRKLFTTENIGCTAYNLGTGRGTSVLEMTATPEKASGKIPVKLC 303

QY 301 RRPDATEVYASTERAELGKWKANYGVEMCRDQWNAKNPWGAGKP 350
 Db 304 RRGDATEVYASTQKAEKELGKWKANYGVEMCRDQWNAKNPWGAGKP 353

RESULT 3
 Q8H057 PRELIMINARY; PRT; 360 AA.
 AC Q8H057;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative UDP-glucose 4-epimerase.
 GN OSUGS-3.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare; TISSUE=Immature seed;
 RA Suzuki K., Kitamura S.;
 RT "Cloning of UDP-glucose 4-epimerase genes in Oryza sativa.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB097460; BAC41499.1; -.
 DR GO; GO:0003824; F:catalytic activity; IEA.

DR GO: 0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO: 0005975; P:carbohydrate metabolism; IEA.
DR GO: 0006012; P:galactose metabolism; IEA.
DR GO: 0009325; P:nucleotide-sugar metabolism; IEA.
DR InterPro: IPR001509; Epimerase_Dh.
DR InterPro: IPR005886; GalE.
DR InterPro: IPR008089; Nuc_sugar_epim.
DR Pfam: PF01370; Epimerase_1.
DR PRINTS: PRO1713; NUCPIMERASE.
DR TIGRFAMs: TIGR01179; GalE; 1.
SQ SEQUENCE 360 AA; 39199 MW; 212BC4D210EA4724 CRC64;

Query Match 72.0%; Score 1328; DB 10; Length 360;
Best Local Similarity 70.0%; Pred. No. 3e-95;
Matches 238; Conservative 50; Mismatches 52; Indels 0; Gaps 0;
QY 8 ILVTGGAGTGHVTVVOLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQFTGD 67
DB 15 VLVTTGAGTGHVTVVIRLEKGFATVVVDNFNSVPEALDRVLLTAGAALRDLFIAGD 74
QY 68 LRNRDLEKLFSTTFDAVHIFAGLKAVASVAKPRRYDFNLTINLYEFMAKYNCK 127
DB 75 LKSDDMKVFAXRAYDAVHIFAGLKAVGESVAHQVYNNVAGTWNLYSAMTKYCK 134
QY 128 MYFSSATVYGQEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAEPEWKILLR 187
DB 135 IVFSSATAYGQEKIPCEDEDFKLQAMPYGRTKLVLENVFRQVQADPEMKVILLRYFD 194
QY 188 PVGAHESGKLGEDPGKIPNNLMFYIQQVAVGRRLTELVNMGHDPYTRDGSALRDYIHVMDL 247
DB 195 P-GAHRSGDIGEDPGKIPNNLMFYIQQVAVGRRLTELVNMGHDPYTRDGTATRDYIHVMDL 254
QY 248 AGCHIAALRKLFTEENIGCTATNLTGRGTSVLEMTAFKASGKKIPVKLCRRPGDAT 307
DB 255 ADCHIAALEKLFATPDIGCVAYNLGTGCGTTTLEVVYKAFESAGKKIPKICRRPGDCT 314
QY 308 EYVASTERAELKLGKANYGVEMCRDOWNAKNPWGYA 347
DB 315 EYVASTDAKELGHSAR:GIEDMCRDOWNAKNPWGYYS 354

RESULT 4
Q8LEA9 PRELIMINARY; PRT; 348 AA.
AC Q8LEA9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE UDP-galactose 4-epimerase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "full-length messenger RNA sequences greatly improve genome
RT annotation."; 0:0-0(2002).
RL Genome Biol. 0:0-0(2002).
[2]
SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: A1085528; AAM62752.1;
DR GO: 0003924; F:catalytic activity; IEA.
DR GO: 0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO: 0005975; P:carbohydrate metabolism; IEA.
DR GO: 0006012; P:galactose metabolism; IEA.

DR GO: 0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro: IPR001509; Epimerase_Dh.
DR InterPro: IPR005886; GalE.
DR InterPro: IPR008089; Nuc_sugar_epim.
DR Pfam: PF01370; Epimerase_1.
DR PRINTS: PRO1713; NUCPIMERASE.
DR TIGRFAMs: TIGR01179; GalE; 1.
SQ SEQUENCE 348 AA; 38312 MW; AFFG6CB99819316 CRC64;
Query Match 66.6%; Score 1229; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 1.5e-87;
Matches 228; Conservative 50; Mismatches 62; Indels 2; Gaps 2;
QY 5 SQHILVTGGAGTGHVTVVOLLKAGFSVSIIDNFNSVMEAVDRVQVVGPLLSQNLQFT 64
DB 2 ARNVLSGGAGYIGSHTVLQLLGGYSVVVDNFNSVLSQVLRVKLAAB-HGERLSFH 60
QY 65 QGDLRRNRDLEKLFSTTFDAVHIFAGLKAVASVAKPRRYDFNLTINLYEFMAKYN 124
DB 61 QVLDLRORSALKFISFKDFDAVHIFAGLKAVGSSVEKPLLYNNLVNLTITLLEVMQAQHG 120
QY 125 CKKRVSSSATVYGQEKIPCEDEDFKLQAMPYGRTKLFLEIARDIQKAEPEWKILLR 184
DB 121 CKLVFSSATVYGSPKEVPECTEEFPISALNPGYRTKLFIEECRDVYGDPEWKILLR 180
QY 185 YFNPVGAHESGKLGEDPGKIPNNLMFYIQQVAVGRRLTELVNMGHDPYTRDGSALRDYIHV 244
DB 181 YFNPVGAHPSGDI GEDPGKIPNNLMFYIQQVAVGRRLTELVNMGHDPYTRDGTATRDYIHV 240
QY 245 MDLADGHIAALRKLFTEENIGCTATNLTGRGTSVLEMTAFKASGKKIPVKLCRRPG 304
DB 241 IDLADGHIAALRKLEDC-K-IGCEVYALGTGNGTSVLEMDVDAFEKASGKKIPLVIAGRRP 299
QY 305 DATEVYASTERAELKLGKANYGVEMCRDOWNAKNPWGY 346
DB 300 DAEVYASTERAELNWKAKYIGIEWCRDLNWNWASNNPYGY 341

RESULT 5
Q8VZ26 PRELIMINARY; PRT; 351 AA.
AC Q8VZ26
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative UDP-galactose 4-epimerase.
GN AT4G10960.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
RA Yanada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At4g10960 (GI:15236988).";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA Yanada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,


```

Query Match      64.0%; Score 1180; DB 10; Length 354;
Best Local Similarity 64.5%; Pred. No. 1e-83;
Matches 227; Conservative 49; Mismatches 72; Indels 4; Gaps 4;

QY 1 MVSS-SQHILVTGGAGFIGHTVTVVQLLKAGFSVSIIDNFDSVMEAVDRVROVWGLLSQ 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MVSAALRTILVTGGAGVIGSHTVLQQLGFRVVLDNLNDSALAILRVRELGH-NAN 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 60 NLOFTGDDLRNRDLEKLSKTTFDVAIHVHAGKAVASVAKPRRYFDNLVGTINLYEF 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 NLDPRKVDLRDQALDQSFSSQRFVAVIHVHAGKAVGESVQKPLLYDNNLIGTITLLQV 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 120 MAKYCKKMYFSSSATVYGOPEKIPCEEDFKLQANMPYGRTKLFLFEEIARDIOKAEPEWK 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 MAHGCCTKLVSSTAVYGMFKEVPCVTEESPLCAMNYPYGRTKLVIEDMCRDLHASDENWK 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 180 IILLRYFNPVGAHSGKLGEDPKGIPNNLMPTIQVAVGRLTELNVYGHDTYTRDGSALR 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 IILLRYFNPVGAHSPGVIGEDPCGIPNNLMPTIQVAVGRRPALTVYGTDTYTKDGTGR 239
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 240 DYIHWMDLADGHTAARLKLFP-TTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKL 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 DYIHWVLDLADGHTAARLKLFP-TTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKL 298
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 299 CPERPGDATEVYASTERAKELGWKANYGVEMECRDOWNWAKNPMGYAGKP 350
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 AGRPGDAEIVYATKAEKELGWKANYGVEMECRDOWNWAKNPMGYAGKP 350
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
ID QH931 PRELIMINARY; PRT; 408 AA.
AC QH931;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative UDP-glucose 4-epimerase.
GN OSUGB-2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare; TISSUE=Immature seed;
RA Suzuki K., Kitamura S.;
RT "Cloning of UDP-glucose 4-epimerase genes in Oryza sativa.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB096863; BAC24803.1; -
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
SQ SEQUENCE 408 AA; 44216 MW; 62D4903994BC7DFC CRC64;

Query Match      62.8%; Score 1159.5; DB 10; Length 408;
Best Local Similarity 64.0%; Pred. No. 4.9e-82;
Matches 217; Conservative 50; Mismatches 71; Indels 1; Gaps 1;

QY 8 ILVTGGAGFIGHTVTVVQLLKAGFSVSIIDNFDSVMEAVDRVROVWGLLSQNLFTQGD 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 14 VLVTGGAGVIGSHAVLQQLLAGFRVAVVLDNLNDSSELAVRVAALAGD-HSNLAFHKVD 72
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 68 LRNRDLEKLSKTTFDVAIHVHAGKAVASVAKPRRYFDNLVGTINLYEFWAKYCKK 127
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 73 LRDKGALKVFASTRDVAHVHAGKAVGESVQKPLLYDNNLIGTITLLQVLSWAGCKK 132
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 128 MVFSSATVYGOPEKIPCEEDFKLQANMPYGRTKLFLFEEIARDIOKAEPEWKIILLYFN 187
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 133 LVFSSAAVYGPSKPNKSPWTEFFPLTNNPYGKTKLVVEDICRDIYRTDPEWKIILLYFN 192
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 188 PVGAHSGKLGEDPKGIPNNLMPTIQVAVGRLTELNVYGHDTYTRDGSALRDIHVMDL 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 193 PVGAHSPGVIGEDPCGIPNNLMPTIQVAVGRRPALTVYGTDTYTKDGTGRDIHVVDL 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 248 ADGHAALRLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCRRPGDAT 307
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 253 ADGHAALQKLFESSIGCEAYNLGTGRTSVLEMTAFKASGKKIPLIIGRRPGDAE 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 308 EYASTERAKELGWKANYGVEMECRDOWNWAKNPMGY 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 313 ILFSLPAKAEKELNWKARFGIDEMCRDOWNWAKNPMGY 351
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
Q9SGX0 PRELIMINARY; PRT; 447 AA.
AC Q9SGX0;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE FIN19.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.;
RA Shinn P., Tambunga G., Altafi H., Bei Q., Chin C., Chiou J., Choi E.;
RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.;
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharisky N., Lam B.;
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.;
RA Thayer A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.;
RA Theologis A., Becker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC FIN19 from chromosome
   I.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009519; AAF19668.1; -
DR HSSP; P09147; IXEL.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; Gale.
DR Pfam; PF01370; Epimerase; 1.
DR TIGRFAMs; TIGR01179; gale; 1.
SQ SEQUENCE 447 AA; 49571 MW; ADC8F8B604BB32D8 CRC64;

Query Match      62.7%; Score 1157.5; DB 10; Length 447;
Best Local Similarity 50.9%; Pred. No. 8e-82;
Matches 225; Conservative 54; Mismatches 64; Indels 99; Gaps 5;

QY 7 HILVTGGAGFIGHTVTVVQLLKAGFSVSIIDNFDSVMEAVDRVROVVG----- 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 NILVTGAGVIGSHTVLQQLLGGYNTVVIDNLNSSLVSLVQKVLADGHNLTVHQVW 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 55 ---PLLS-----QNLOFTQG----- 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 LSPFLCSDLCHILRLISVIMTSKIESFFVFRFAYNALGMVFQIAFRSGFLDQIVFEI 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 67 --DLNRDDLEKLSKTTFDVAIHVHAGKAVASVAKPRRYFDNLVGTINLYEFWAKYN 124
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 IVDLRDPALEKVFSETKFDVAVMEFAGLKAVGSSVAKPLLYNNNLIAITITLLEWMAAG 183
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 125 CKKWVSSSATVYGOPEKIPCEEDFKLQANMPYGRTKLFLFEEIARDIOKAEPEWKILLR 184
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 CKKLVSFSSSATVYGNPKEVCTEESPLSGMSPYGRTKLFTIEDICRDVQRGDPEWRIIMLR 243
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
QY 185 YFNPVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGS----- 237
Db 244 YFNPVGAHESGRIGEDPCGTPNNLMPYVQVVGRLFNKLYIGDYTTKDGTVRLSHPY 303
QY 238 -----IRDYHVMDLADGHHIAALRLKLTFTENIGCTA 268
Db 304 TTRINLNCPLNPCIITLTCTCFLSFHNNHVDYIHVVLDLADGHICALQKLDITE-IGCEV 362
QY 269 YNLGTGRTSVLEMTAFKASGKKIPVKLCPRPGDATFYASTERAELGKWKNYGV 328
Db 363 YNLGTGRTGTVLEMTAFKASGKKIPVKLVGRPGDAETVYASTERAELNWKANFGI 422
QY 329 EEMCRDOWNWAKNPNWGVAGKP 350
Db 423 EEMCRDOWNWASNPNFGYGSP 444

RESULT 12
Q81414
ID Q81414 PRELIMINARY; PRT; 338 AA.
AC Q81414;
DT 01-JUN-2003 (TremBrel. 24, Created)
DT 01-JUN-2003 (TremBrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2).
GN BC5448.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapatral V., Bhaktacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Coleman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haseelkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrgides N.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis."
RL Nature 423:87-91 (2003).
DR EMBL; AE017015; AAP12309.1; -.
DR GO; GO:0016853; P:isomerase activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GALE.
DR InterPro; IPR00205; NAD_BS.
DR InterPro; IPR008089; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 338 AA; 37270 MW; 93225A3A6ACA3DD4 CRC64;
```

```
Query Match 56.7%; Score 1046; DB 16; Length 338;
Best Local Similarity 56.6%; Pred. No. 2.6e-73;
Matches 192; Conservative 58; Mismatches 83; Indels 6; Gaps 2;

QY 8 ILVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDNSVMEAVDRVQVGVPLLSQNLQFTQGD 67
Db 3 ILVTGGAGYIGSHTCVELLNSGVEVLVDNLSNSVESINRVKHTG----KQFXYKED 58

QY 68 LRNRDDEKLFSTTTDAVTHFAGLKAVASVAKPRRYDFNLVGTINLYEFMAKYNCKK 127
Db 59 VLNRALDAIFBENAIEAVTHFAGFVAVGESVAIPITYHNNTSLVLCEVWQXNVCK 118

QY 128 WVFSSATVYGOEKKIPCEEDFKLQANPNYGRTKLFLEIARDIQKAEPEWKIILIRYFN 187
Db 119 IFSSATVYGIPTETPTEEPFLSVTPNYGQTKLMIEQIMRDVAPADAENSIALLRYFN 178
```

```
QY 188 PVGAHESKLGEDPKGIPNNLMPYIQQVAVGRLTELNVYGHDPYTRDGSARDYIHVMDL 247
Db 179 PVGAHESGRIGEDPCGTPNNLMPYVQVVGRLFNKLYIGDYTTKDGTVRDIHVVDL 238
QY 248 ADGHIAALRLKLTFTENIGCTAVNLGTGRTSVLEMTAFKASGKKIPVKLCPRPGDAT 307
Db 239 ANGHVAKLEKVL--RTGVDAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRPGDA 296
QY 308 EYVASTERAEKLGKANYGVVEEMCRDOWNWAKNPNWGY 346
Db 297 VCFADASKAKRELGWATRGLEEMCADSWKQSNKNGY 335

RESULT 13
Q81JK4
ID Q81JK4 PRELIMINARY; PRT; 338 AA.
AC Q81JK4;
DT 01-JUN-2003 (TremBrel. 24, Created)
DT 01-JUN-2003 (TremBrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE UDP-glucose 4-epimerase.
GN GALE-2 OR BA5700.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad C.A., Helgason E., Rilstone G., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman J.F.,
RA Hazen A., Cline R., Redmond C., Thwaiter J.E., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria."
RL Nature 423:81-86 (2003).
DR EMBL; AE017041; AAP29332.1; -.
DR TIGR; BA5700; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
DR InterPro; IPR001509; Epimerase_Dh.
DR InterPro; IPR005886; GALE.
DR InterPro; IPR008089; Nuc_sugar_epim.
DR Pfam; PF01370; Epimerase; 1.
DR PRINTS; PR01713; NUCEPIMERASE.
DR TIGRFAMs; TIGR01179; gale; 1.
KW Complete proteome.
SQ SEQUENCE 338 AA; 37845 MW; F24E14DB2B62CF9 CRC64;
```

```
Query Match 55.4%; Score 1023; DB 16; Length 338;
Best Local Similarity 54.3%; Pred. No. 1.6e-71;
Matches 184; Conservative 63; Mismatches 86; Indels 6; Gaps 2;

QY 8 ILVTGGAGFTGTHVTVOLLKAGFSVSIIDNFDNSVMEAVDRVQVGVPLLSQNLQFTQGD 67
Db 3 ILVTGGAGYIGSHTCVELLANNYKIIIVVDNLSNSESINRVKHTG---KQFEFYKEN 58

QY 68 LRNRDDEKLFSTTTDAVTHFAGLKAVASVAKPREYDFNLVGTINLYEFMAKYNCKK 127
Db 59 VLNRKNEIFLENNIEAVTHFAGFVAVGESTTTPLAYTYNNIIISAIVLQVWQXNVCK 118

QY 128 WVFSSATVYGOEKKIPCEEDFKLQANPNYGRTKLFLEIARDIQKAEPEWKIILIRYFN 187
Db 119 FIFSSATVYGIPTKLTPEEPFLSVTPNYGQTKLMIEQIMRDVAKADDEWSIALLRYFN 178
```

QY 188 PVGAHESGKLGEDPKGIPNNLMPIYQVAVGRITELNVYGHDPTRDGSAIRDYIHMDL 247
 Db 179 PFGAHSQGRIGEDNGIPNNLMPIYQVAVGRITELNVYGHDPTRDGSAIRDYIHMDL 238
 QY 248 ADGHIAALRLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRRPDAT 307
 Db 239 AKGHVKALEKVLTK--GIEAYNLGTGRTSVLEMTAFKASGKKIPVKLCPRRRPDAT 296
 QY 308 EVYASTERAELGKWKANYGVENECRQDQWAKNPNWGY 346
 Db 297 ICFADVSKAKRELGWAEYGLEENCVDMSRWQVNNKNGY 335

RESULT 14

Q7ZZM6 PRELIMINARY; PRT; 348 AA.
 AC Q7ZZM6; MEDLINE=22388257; PubMed=12477932;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan A., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX Klein S., Strausberg R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051601; AAH51601.1; -;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.

DR InterPro; IPR008089; Nuc_sugar_epim.
 DR Pfam; PF01370; Epimerase; 1.
 DR PRINTS; PRO1713; NUCPEIMERASE.
 DR TIGRFAM; TIGR01179; galE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 348 AA; 38257 MW; 302A809C103AC173 CRC64;
 Query Match 54.2%; Score 999.5; DB 13; Length 348;
 Best Local Similarity 55.9%; Pred. No. 1.2e-69;
 Matches 194; Conservative 50; Mismatches 90; Indels 13; Gaps 4;
 QY 8 ILVTGGAGFIGTHTVOLLKAGFSVSIIDNPDNSVM-----EAVDRVQVVGPLLSQNL 61
 Db 5 VLVTGGGIGSHCVLELEAGYSPVVIDNFHNAIRGANNIPESLHRVHDIVG---KTT 60
 QY 62 QFTQGLNRDDLEKLFSTKTTTDAVIHFAGLKAVASVAKPRRYFDNLTINLYEFMA 121
 Db 61 EFEEVDILDRALDKLFKSKHFSVAVLHFAGLKAVGESVQKPLLYKVNLTGTIQLLEVN 120
 QY 122 KYNCKWVSSSATVYGOPEKIPCEEDFKL-QANNPYGRTKLFLEETARDIQKAEPEWKI 180
 Db 121 SHGVNIVFSSSATVYGDNPYLPIDESHFVGGCTNPYKTKYFIEEMIKDOCKAEEDWCA 180
 QY 181 ILLRYFPVGAHESGKLGEDPKGIPNNLMPIYQVAVGRITELNVYGHDPTRDGSAIRD 240
 Db 181 ILLRYFPVGAHESGKLGEDPKGIPNNLMPIYQVAVGRITELNVYGHDPTRDGSAIRD 240
 QY 241 YIHVMDLADGHIAALRLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLC 300
 Db 241 YIHVMDLADGHIAALRLFTTENIGCTAYNLGTGRTSVLEMTAFKASGKKIPVKLC 300
 QY 301 RRPDATEVTVASTERAELGKWKANYGVENECRQDQWAKNPNWGY 347
 Db 299 RREGDIATCYADPALAKAELGWAEYGLDRCEDLNRWQACNPPTGFS 345

RESULT 15

Q8EGE0 PRELIMINARY; PRT; 337 AA.
 AC Q8EGE0; MEDLINE=22297686; PubMed=12368813;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE UDP-glucose 4-epimerase.
 GN GALE OR S01664.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RX Heidelberg J.F., Paulsen I.T., Neilson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Sehadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Manuvelan J., Weidman J., Imprim L.M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
 RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RT Shewanella oneidensis";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 DR EMBL; AE015612; AAN54719.1; -;
 DR TIGR; S01664; -;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0003978; F:UDP-glucose 4-epimerase activity; IEA.
 DR GO; GO:0006012; P:galactose metabolism; IEA.
 DR GO; GO:0009225; P:nucleotide-sugar metabolism; IEA.
 DR InterPro; IPR001509; Epimerase_Dh.
 DR InterPro; IPR005886; GalE.
 DR InterPro; IPR000205; NAD_BS.

